

DETERMINING MEAN GRAIN-SIZE IN HIGH GRADIENT STREAMS WITH
AUTOCORRELATIVE DIGITAL IMAGE PROCESSING

A Thesis

by

CARLA ANN PENDERS

Submitted to the Graduate School

Appalachian State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

May 2010

Department of Physics and Astronomy

Copyright by Carla Ann Penders 2010
All Rights Reserved

ABSTRACT

DETERMINING MEAN GRAIN-SIZE IN HIGH GRADIENT STREAMS WITH AUTOCORRELATIVE DIGITAL IMAGE PROCESSING

(May 2010)

Carla Ann Penders, B.S., Appalachian State University

M. A., Appalachian State University

Chairperson: Christopher S. Thaxton

Accurate surface roughness characterization of a stream bed is essential to hydrodynamic and sediment transport modeling. High gradient streams are particularly challenging in part because small changes in the mean grain size can have a demonstrative impact on its fluid dynamics. Traditionally employed methods of determining mean grain size are labor intensive, provide low temporal resolution, and are historically inaccurate. We present an automated digital image processing method for high gradient streams that uses spatial autocorrelation to determine mean grain size. Numerical results are compared to statistics obtained from sieve-based protocols. Results support this method as a potential alternative to traditional methods with a dramatic increase in temporal resolution. When optimized parameters are applied the system accurately determines mean grain size within a 50% difference from sieved samples 85% of the time.

ACKNOWLEDGEMENTS

This research was supported by the Office of Student Research at Appalachian State University in 2007 and most substantially by the North Carolina Space Grant in 2007, 2008, and 2009.

I also wish to acknowledge the support of my thesis committee members Dr. Richard Gray, Dr. Roy Sidle and Dr. Tonya Coffey. I wish to express special thanks to my committee chairperson and mentor Dr. Christopher Thaxton without whose continuous support and encouragement this research would not have been completed.

DEDICATION

This work is dedicated to my three amazing children: Kaitlyn, Kelson, and Jaden. You are the light of my life and the main reason I set out upon this academic journey. It is my fervent hope you will consider my example to be one worth following.

TABLE OF CONTENTS

Abstract.....	iv
Dedication.....	v
Acknowledgements.....	vi
Chapter 1: Introduction.....	1
Chapter 2: Methods.....	6
Data Acquisition.....	6
Image Processing.....	10
Parameter Optimization.....	15
Error Quantifictaion.....	18
Chapter 3: Data and Results.....	20
Data.....	20
Optimization Results.....	22
Results Under Optimized Conditions.....	33
Chapter 4: Conclusions and Recommendation.....	36
Works Cited.....	42
Appendix A (Tables 7 Graphs).....	45
Appendix B (Images).....	58
Biographical Sketch.....	65

INTRODUCTION

The accurate characterization of granular roughness in a stream bed is essential for hydrodynamic and sediment transport modeling and is of great importance for engineering work in the area of sedimentation. In the absence of bedforms and for sheet flow conditions, granular roughness solely determines the friction coupling between the fluid and the bed, and is normally parameterized through a representative grain size such as, for example, median grain diameter (D_{50}), mathematical mean diameter (D_m), or geometric mean diameter (D_{gm}) (e.g., Haan et al. 1994). This parameterization may be used in conjunction with distribution functions such as a log-normal probability density function or cumulative grain size curve (e.g., Fredsoe and Deigaard 1992; Nielsen 1992). Representative grain size is also a fundamental variable of constitutive relations such as those governing settling velocity, drag force, critical shear stress, and Shields and Reynolds numbers, (e.g., Fredsoe and Deigaard, 1992; Nielsen 1992). For mountain streams, small grains and wash load are selectively entrained due to high gradient-driven shear stresses (Gasparini et al. 1999). Small grains scour about the larger grains (e.g., Nino et al. 2003) (see figure 1) and are transported away, leaving only large grains (gravel, cobbles, boulders) with interstitial and patchy areas dominated by sand and silt ($>0.05\text{mm}$, USDA 1979). In many mountain streams, “granular roughness” becomes less descriptive of the boundary layer dynamics – the “grains” act more like bedforms in terms of their extended vertical effect on the turbulent structure within the water column, effectively thickening the boundary layer (see figure 2).



Figure 1. In mountain streams, large grains dominate (left) in which small grains (silt and clay) are preferentially transported downstream, for example to the Mississippi River delta (right). From www.zazzle.com (left) and www.stchas.edu (right).

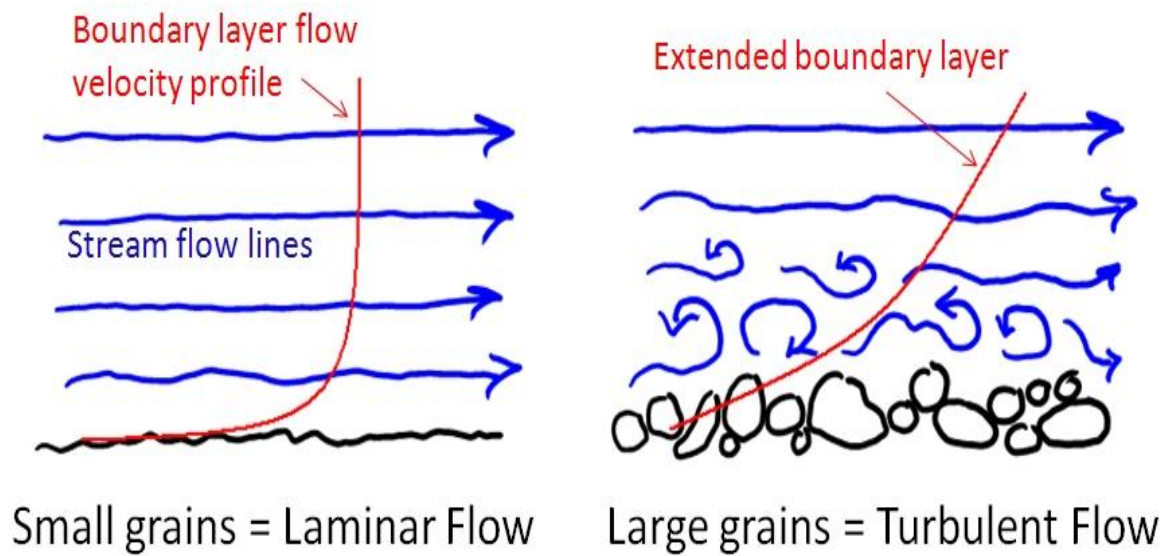


Figure 2. Sketch of the relative behavior of stream flow for a shallow boundary layer due to small grains (left) and a thick boundary layer due to large grains.

Thus, accurate characterization of the spatial patchiness of grains and temporal variations in a stream's bed-surface grain size distribution is critical for accurate hydrodynamics modeling, arguably more so than for river deltas, estuaries, and littoral zones.

Grain size determination has historically been done using a variety of methods including direct measurements (ruler, caliper, etc.), sieving, elutriation, sedimentation,

permeability, etc. (Loveland and Whalley 2001). For standard sieving methods, which are the most prevalent and least expensive, grab samples are obtained from multiple cross-stream sites and dried and sieved according to well-documented protocols (ASTM D421-85 2007). A significant amount of effort and time is required to characterize relatively small reaches of a stream using standard sieving methods. As a result, grain size characterization is normally done only occasionally which makes it difficult to frequently and reliably characterize stream beds in support of sediment transport and bed dynamics analysis and modeling. In addition, manual methods are historically inaccurate (e.g., Gluschke et al. 2004) due in part to shape and density variations among the grains. Review of current literature suggests that grain size estimation from standard methods includes error on the order of 200% or more (Penders and Thaxton, 2010)

Multiple automated photometric methods have been developed to address the stated need for grain sizing methods with lower time investments and higher accuracies (e.g., Adams 1979; Ibbeken and Schleyer 1984). Many disciplines employ a wide range of currently evolving edge-detection algorithms, ranging from cosmic ray rejection (van Dokkum 2001) to arterial wall tracking (Woodman et al. 2001) to auditory neurological studies (Fishbach et al. 2001) and beyond, with improvements over conventional methods. However, even recent, advanced computational routines (Meer & Georgescu 2001) struggle to resolve “weak edges” characteristic of low contrast gradients commonly found in images of stream bed sediments. Sime and Ferguson (2003) used edge detection methods for determining grain size distributions from images at low resolution. Butler et al. (2001) complemented binary edge detection with “edge growth” to resolve individual grains more clearly. Zhou et al. (2004) employed a “seeded region-growing” edge detection and

resolution algorithm to define grains in color images. During initial research on the project described herein, we attempted edge-detection using the built-in ERDAS algorithms, including 1st-order derivative, Zero-sum filters, and 2nd-order derivative methods (ERDAS, 1999) with minimal success. In general, most automated grain sizing techniques researched to date are either marginally effective (e.g., McEwan et al. 2000) or are too computationally intensive for fast, field-based analysis, or both.

Dave Rubin (2004) developed an autocorrelation routine that has been used successfully in limited applications, mostly in dry beds. Rubin employed an autocorrelation algorithm to an image, treated as an N x M matrix of pixel intensities, in which a “data” image of mixed sediment was decomposed via non-negative least squares into its constituent eigenstates within a basis set of “calibration curves” obtained from a set of images of fixed grain size. Others (Rubin, et al., 2007; Buscombe 2008; Buscombe and Gerhard, 2009) extended Rubin’s work to characterize a broader set of statistical measures. Rubin’s method freed researchers from the computationally intensive and inherently problematic edge detection approaches to grain size characterization while showing, at least for dry beds, a consistent and marked improvement over standard methods.

Herein, we describe a method based on Rubin (2004) in which we apply the autocorrelation routine to permanently wetted mountain streams. Numerical results are compared to statistics obtained from sieve-based protocols. Data were obtained over a range of control variables, including photometric variables (magnification), image processing variables (rotation, inversion, segmentation, format), and numerical variables (autocorrelation offset). Trend and sensitivity analysis were done on the full range of variables for the grain size limits applicable to mountain streams in the upper New River

Watershed, North Carolina, USA. Results indicate that, within a tolerance of 50% difference from traditional methods, the mean grain size is accurately determined 85% of the time. The system was robust for compressed images. The speed with which data may be collected and processed makes application of this method preferable due to the dramatic increase in temporal resolution in support of stream modeling

METHODS

Data Acquisition

Photographs of stream bed surface sediment were obtained from two streams: a section of Boone Creek that runs through the campus of Appalachian State University (summer of 2007), and a portion of the New River that runs through the Greenway (summer of 2008) (see figure 3) both of which are located in Boone, NC USA. Photographs were also collected in the lab for system calibration. All images were collected using an Olympus SP-510UZ, 7 megapixel digital SLR camera.

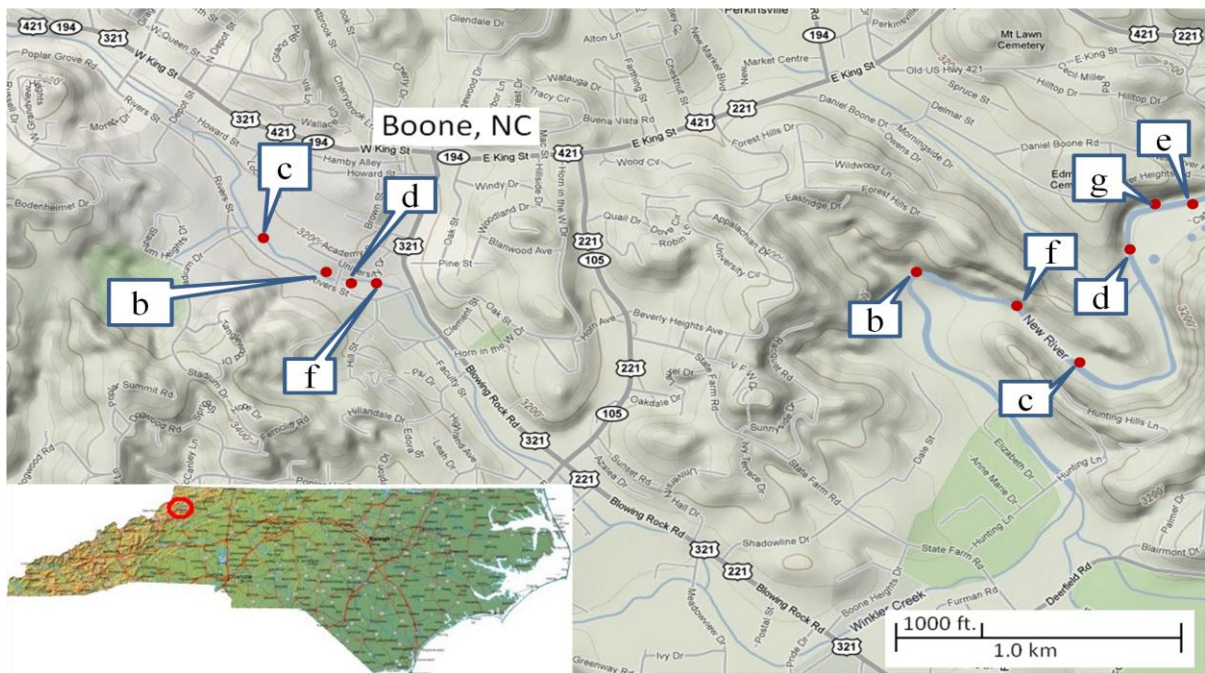


Figure 3. Data was gathered at four locations in Boone Creek (BC) located on the left side of the map and at six locations along the New River (NR) seen on the right side. Each location identifier can be found in the image identification code along with other pertinent information. (Site (a.) is not shown – data was not included due to non viable images)

An aluminum frame was designed to maintain constant distance of 1.0 m and a 90° angle between the camera-lens center line and the surface of the sediment surface (See schematic in figure 2-A). The base provides a 1m x 1m square image area. The camera attaches securely to the mount for easy maneuvering in the field and lab. The frame breaks down to the base piece, the camera mount and four legs which can be partially or completely detached if needed for portability. In order to minimize optical distortion of the image caused by surface ripples in moving water, a PlexiglasTM box was created to mimic a ‘glass bottom boat’ effectively flattening the water and removing the distortion. It was made of one quarter inch PlexiglasTM with dimensions: 24”x18”x12”. A small square rod (¼ “ x ¼ “ x 12”) of the same material was placed on the inside of each corner for stability. Clear silicone caulk was used to seal the container. A specialized cleaner made for small airplane windshields, Plexi-ClearTM (Jet Stream Aviation Products, Dallas, TX.) was used to keep the bottom surface clean and free of air bubbles when submerged. Keeping this surface free of debris or any scratches is critically important for collecting viable data.

In the field, two pieces of rebar about 1.3m in length were driven approximately 0.3m deep into the streambed about 0.5m apart to maintain the position of the PlexiglasTM box in the stream. (See figure 4 and appendix B) This served to prevent the box from moving out of the image area, particularly in areas of swiftly moving water. A large black umbrella was used to provide shade for the image area. This was necessary to remove reflection of light from the inside surface of the PlexiglassTM box, which could alter the image so that the grains were not visible in the image. Photographs were generally taken in the early morning, evening or on cloudy days.

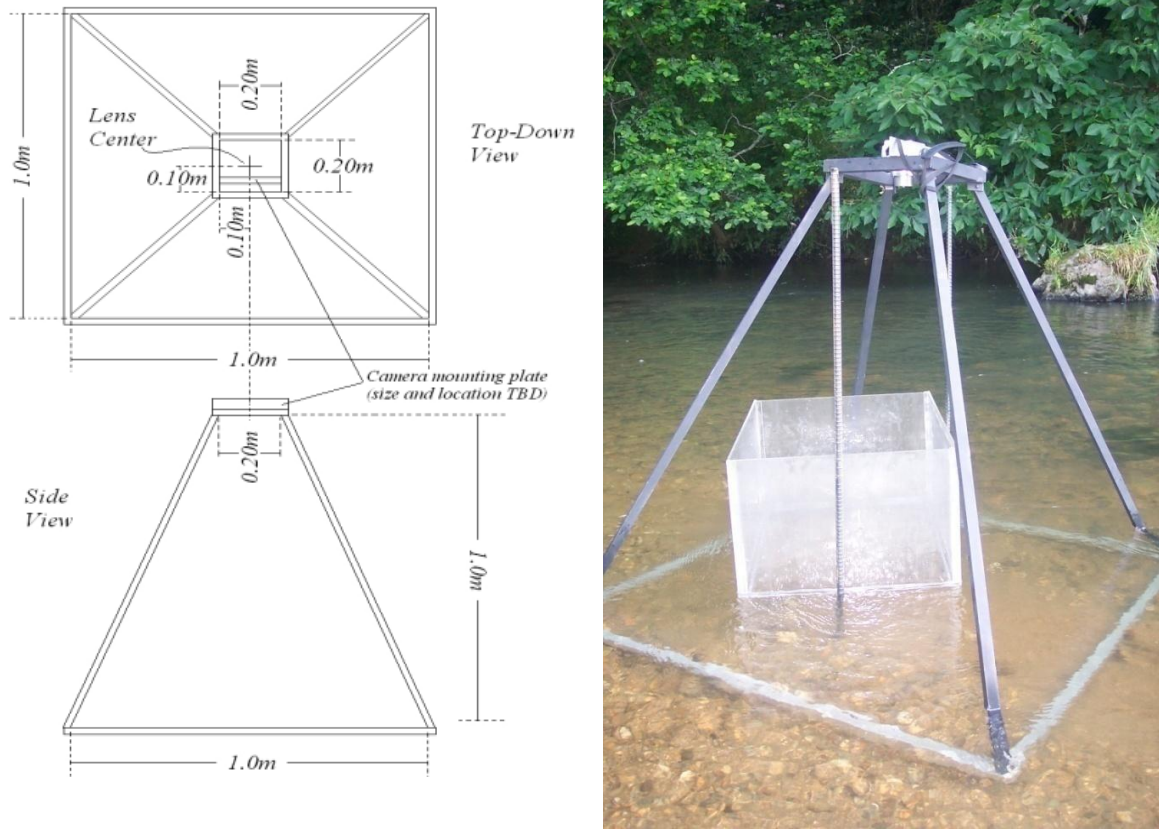


Figure 4. On the left, a schematic of the aluminum frame used to fix the position of the camera 1 meter from the stream bed and at a 90o angle. On the right is the apparatus set up for data acquisition in the New River.

The camera provided a variety of resolutions via focal-length (magnification) settings. At each magnification, the number of pixels per millimeter was determined manually. During processing, we determined that a minimum of 7 pixels were required to effectively and repeatably identify unique grains, which forced a base value of 10 for the lowest maximum pixel offset, discussed below. See table 1 for resolution details.

Herein, the 50mm, 100mm, and 200mm equivalent focal length settings were used exclusively, with 200mm being the most commonly utilized. All data images were taken without a flash, both in the lab and in the field to maintain consistency.

Table 1. Camera Magnification/Resolution details including the number of pixels per millimeter, minimum grain size resolved and image dimension

Focal Length (magnification)	Pixels/cm	Pixels/mm	Minimum Grain Size Resolved (mm)	Image Dimension (cm)
38 mm (min)	35	3.5	2 mm	90 x 67
50 mm	45	4.5	1.6 mm	70 x 52
100 mm	91	9.1	0.8 mm	32 x 25
200 mm	166	16.6	0.44 mm	19 x 13
380 mm (max)	265	26.5	0.27 mm	12 x 9

Sediment sample collection was performed according to traditionally accepted protocols (ASTM D421-85 2007) with an emphasis placed upon collecting sediment nearest to the surface of the streambed. Immediately following image acquisition in the field, a sediment sample was collected, placed in a plastic bag, and marked according to its location and position in the body of water it was taken from. A corresponding note in a field log identified the picture number for that day, along with the location and position in the water, as well as the magnification that was used. Thereafter the sediment was dried at approximately 100 °C for no less than 24 hours. Once dried it was sieved using a shaker (Humboldt Manufacturing, Schiller Park, IL.) with 10 standard sieves (Fisher Scientific, Pittsburgh, PA) sized: 0.25mm, 0.50mm, 1mm, 2mm, 4.7mm, 8mm, 16mm, 25mm, 50mm, 100 mm. The contents retained in each sieve were weighed to the nearest 0.02 grams using a triple beam balance scale (OHAUS, Pine Brook, NJ). The mathematical mean grain size was then determined and recorded for each sample using standard calculations.

Image Processing

Each image was prepared for processing using Corel Paint Shop Pro™. Any images out of focus, with edges outside the image field defined by the Plexiglas™ box, anomalous reflections, scratches, debris or fish floating through the image field were excluded from the final data set. All useful images were cropped to a standard size of 3100 pixels by 2800 pixels if needed and saved as a tagged image format file (.TIF).

A log of each image was kept to maintain a direct link between the original image names generated by the camera software and the processed image names used for identification thereafter. Each image was renamed with a code that included information about the date of acquisition, body of water it was taken from, the position in the stream, and the magnification used. Consider, for example, the identifier: NR625B01A2. NR is for the New River, 625 refers to June 25th, B refers to the site in the river (see figure 1), 01 refers to the image number of that location, A refers to the cross sectional location, and the last number, 2, refers to the magnification (200 mm in this case). The cross sectional location describes one of four or five sample locations at that particular site. For example, at site B in the New River, five different locations were sampled that follow a straight line perpendicular to the stream flow direction – this was done to get an idea of how the mean grain size changes from one side of the stream to the other. A Dell, Inc. Inspiron 6000 laptop, and then later a Dell, Inc. Precision T7400 workstation were used to store and prepare images, as well as run the Matlab™ code to process them.

Each image was converted to grayscale in which pixel intensity values ranged from 0 to 255. A modified version of an autocorrelation algorithm (Rubin 2004), run in MATLAB™, was developed and employed which compared consecutive pixel intensity values at predetermined maximum offsets. This was done for both calibration images as well as data images. A detailed description follows.

Calibration curves were developed for each sieved grain size from images of pre-collected and dried sediment samples obtained from the streams used in this study. Once the sediment was sieved into the 10 chosen sizes (0.25mm – 100mm as discussed above), images were taken of each sample in the lab. Each image was then processed using the autocorrelation algorithm which resulted in a unique set of correlation coefficients, or a “calibration curve,” for each size. The complete set of calibration curves, one for each unique grain size, constituted a basis set of eigenvectors used later for eigenvalue decomposition of “data” images of mixed grains to obtain grain size statistics.

The autocorrelation routine compares two identical matrices using the simple 2 dimensional correlation coefficient equation:

$$r = \frac{\sum_i \sum_j A_{ij} - \bar{A} \quad B_{ij} - \bar{B}}{\sqrt{\left(\sum_i \sum_j A_{ij} - \bar{A}^2 \right) \left(\sum_i \sum_j B_{ij} - \bar{B}^2 \right)}} \quad (1)$$

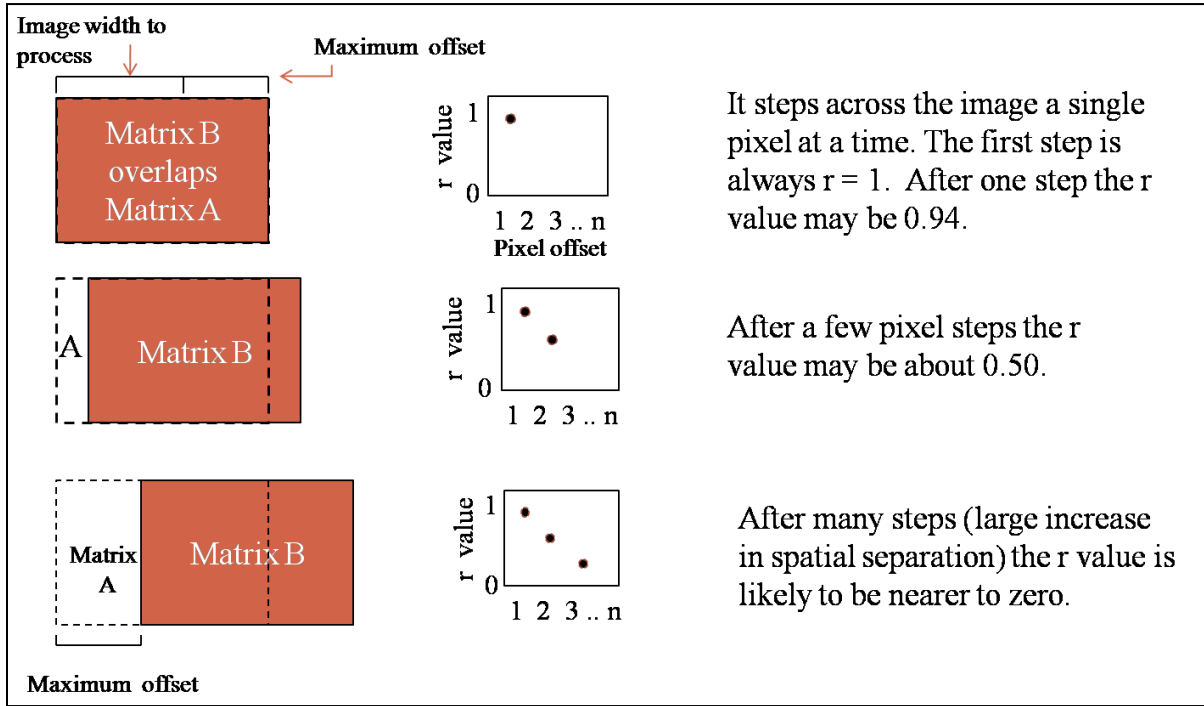


Figure 5. The autocorrelation begins with both matrices overlapping exactly. Matrix B then ‘steps’ one pixel at a time across the image, determining the correlation coefficient at each step. Each of these steps results in a point. After many steps these points represent a curve of coefficient values.

Where A_{ij} and B_{ij} are individual matrix elements (pixels). \bar{A} and \bar{B} are the average values for each respective matrix. For our purposes, matrix A and matrix B are identical matrices made up of the gray scaled pixel light intensity values in a single image of sediment. Matrix B is stepped across matrix A one pixel at a time (see figure 5). Generally speaking, pixels are decreasingly correlated as one moves away from the start point and in a pattern that is determined by the number of pixels per grain. This process steps through the image width to process which is limited by the maximum offset and results in a set of ‘ r ’ values that graph as a ‘correlation curve’ that is unique to any particular grain size (see figures 5 and 6).

Since the calibration curves are dependent on the autocorrelation maximum offset value, a range of offsets for each calibration set was created (10 to 800 pixels to be discussed

below). This allowed data images of mixed grain sizes to be compared to calibration curves at a variety of maximum offsets which, in turn, allowed for the possibility of an optimal maximum offset determination. Figure 6 shows a set of calibration curves for the maximum offset of 50 pixels (camera magnification 200 mm). The curves that approach zero most quickly correspond to the images of small grain sizes while the ones that remain correlated for longer correspond to the images of large grains.

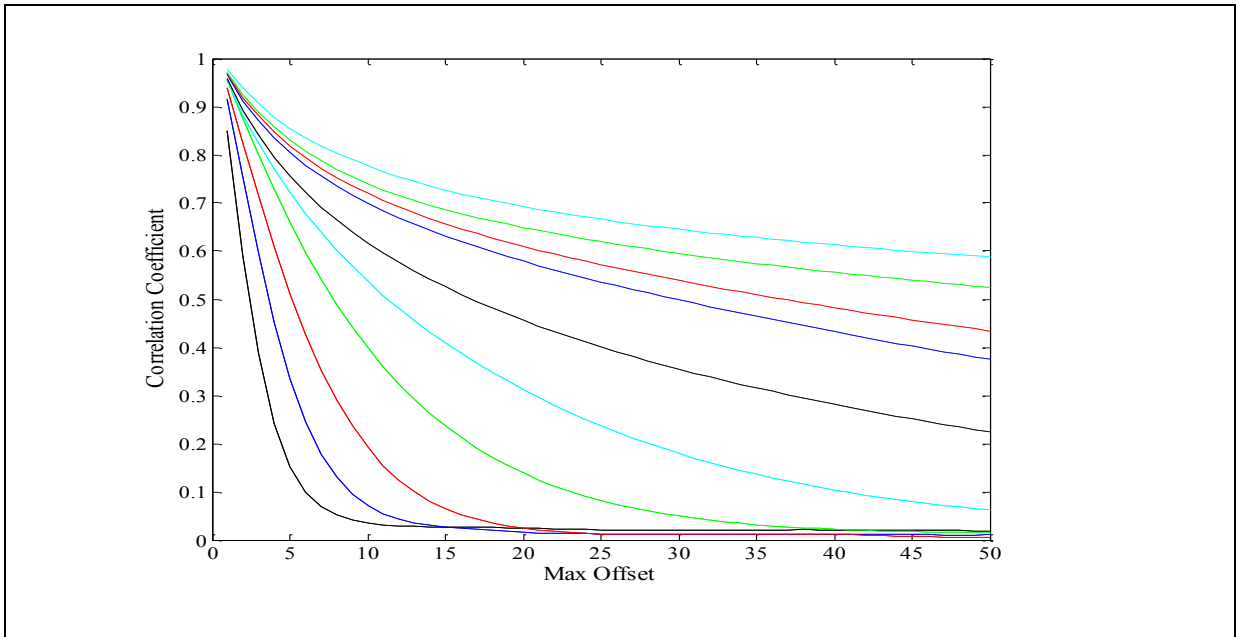


Figure 6. A plot of all ten calibration curves for focal length 200 mm and a maximum pixel offset of 50. A data image curve generated for the same focal length and maximum offset is then compared to this set of calibration curves to determine mean size.

Rotation and transposition of the calibration and data images was implemented to refine the accuracy of the system, particularly in the larger grain size range. Each image is input as a matrix of pixel intensity values. Rotation of the values increases the amount of the image processed, because the image-width-to-process is actually a strip along one side of the image. Rotating four times by 90 degrees allows for all four sides of the image to be processed. Transposing the matrix values results in a mirror image of the original set which

allows the algorithm to run horizontally *and* vertically through the image width to process without disturbing the integrity of the grains. This led to a more effective representation of oblong and irregularly shaped grains in the sample. A total of eight orientations were processed for each image with the results averaged to produce a unique curve of r values.

Segmentation of images was used to accommodate the high variability in grain sizes and grouping of like sized grains within an individual image (see figure 7). Each image was split into segments with each segment being processed individually as discussed. The results were then averaged with those of the other segments in the same image to get a single r value per image. A single segment (the whole image), 4 segments, 9 segments, and 16 segments were all run for both calibration images as well as data images to facilitate defining an optimal number of segments that should be used in general.

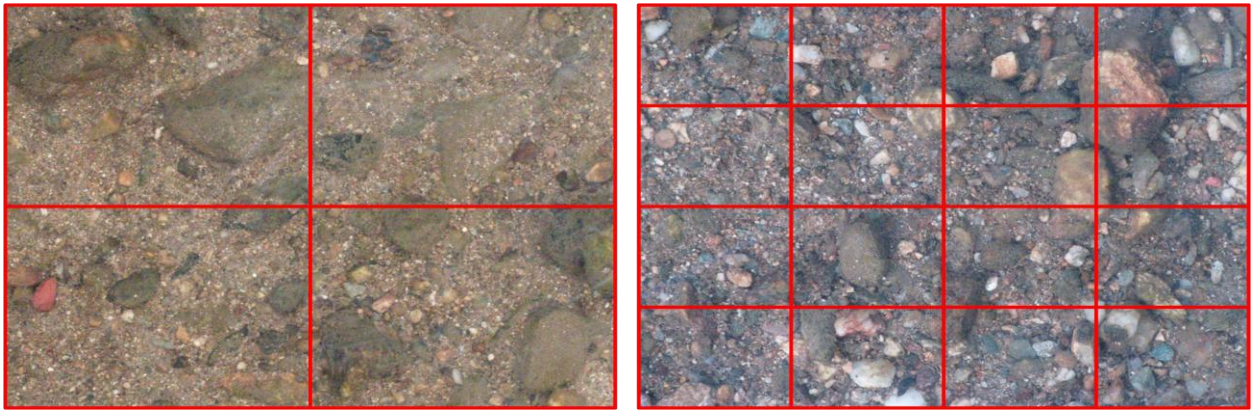


Figure 7. Example of how images are split by the algorithm. Each segment is run separately then the results are averaged to create a curve unique to that image. The segments can be 1, 4, 9, or 16.

The algorithm decomposes the data curve into an eigenvalue representation using a non-negative least squares fit routine, with each normalized eigenvalue representing the percent composition for each calibrated grain size. This distribution is then used to determine the mean grain size of the data image using the simple mathematical mean.

Parameter Optimization

Per the data acquisition and image processing methods described above, multiple control parameters exist for the autocorrelation grain sizing method including acquisition parameters of magnification (focal length) and distance-to-surface (here, fixed at 1.0m) as well as processing parameters for both calibration and data images of maximum offset, number of orientations (rotations and transpositions), and number of image segments. As a result, we sought to define an optimal range of parameters for which the autocorrelation method best matched the sieving results for mathematical mean grain size.

In the lab, 60 images of sediment were obtained ranging in mean size from 3mm to 19mm which were used to determine optimized values for camera magnification, data image segmentation, maximum offset, and calibration segmentation. 20 samples were imaged at 50mm, 100mm, and 200mm focal lengths (equivalent camera magnification). The camera was not moved, nor the sediment disturbed between magnifications, to preserve as much consistency as possible. The images for each magnification were then run through the autocorrelation algorithm at maximum offsets ranging from 200 pixels to 800 pixels. Each image was run at 1, 4, 9, and 16 segments for each of 5 maximum offsets.

. The sediment used for optimizing parameterization consisted of mixtures of dry sieved sediment from the New River and Boone Creek. Each mixture's mean grain size was determined using accepted sieving protocols.

Analysis required approximately 50 code runs for each magnification, each signifying a different combination of computational variables. Results were then put into a group of five matrices for each magnification, one matrix for each maximum pixel offset used. Resolution is one factor that limits the maximum offset values that can be used as well as the number of segments that an image can be effectively split into. Segmentation also limits the offsets available. This means that none of the matrices have all possible values and that different magnifications used overlapping ranges of offsets to maximize the number of entries. (See Table 2 which describes the pixel offset values used for each magnification matrix.) The minimum value for the maximum offset was determined initially to be one that included the largest *mean* grain size of the lab samples imaged. For example, a 20 mm grain at a resolution of 16 pixels per mm requires 320 pixels. Further discussion of maximum offset will be presented in the results section.

Table 2. Maximum pixel offsets associated with magnification matrices.

50 mm Magnification	100 mm Magnification	200 mm Magnification
200 pixel maximum offset		
300 pixel maximum offset	300 pixel maximum offset	
400 pixel maximum offset	400 pixel maximum offset	400 pixel maximum offset
500 pixel maximum offset	500 pixel maximum offset	500 pixel maximum offset
600 pixel maximum offset	600 pixel maximum offset	600 pixel maximum offset
	700 pixel maximum offset	700 pixel maximum offset
		800 pixel maximum offset

Matrix axes consisted of data image segments and calibration image segments (1, 4, 9, and 16 respectively) Each value in the matrix represents an evaluation of accuracy for a specific variable set. Accuracy of an individual image mean size estimate was determined as

a percent difference from traditional sieved mean grain size. Each matrix value corresponds to the percentage with which the data set was accurately assessed. For example, a 75 as one entry means that for that particular combination of variables the mean grain size was accurately determined for 75% of the lab images within specified tolerances. See Table 3 and the Appendix A.

Table 3. Example matrix for 100 mm focal length and 400 pixel offset. Along the top are the number of segments for the data images, down the left side are segments for the calibration images. There were 5 matrices for each camera resolution, one for each maximum offset used. Each value is the result of a single code run for that particular set of variables.

	1	4	9	16
1	35	45	60	70
4	55	60	60	65
9	20	25	40	30
16	10	5	x	x

The x's in table 3 indicate that for these images taken at 100 mm magnification data image size restricted the segmentation to below 9 segments. The entries on the table show that for a single segment calibration image, a sixteen segment data image, at a 400 pixel offset and 100 mm magnification, the estimated mean grain size is within 25% difference of the traditionally-determined mean grain size 70% of the time.

Tolerances were chosen to be any estimated value within a specified percent difference range. Five tolerances were chosen to illustrate system effectiveness: 15%, 25%, 50%, 75%, and 100% difference from the traditionally determined mean grain size. A complete list of the matrices generated for this research can be found in Appendix A.

Error Quantification

Accurately quantifying error for this system is problematic. Clearly there is error inherent to the traditional methods currently used to determine mean grain size and there is also error associated with digital processing and use of the autocorrelation algorithm. In addition, neither traditional error analysis nor digital processing error analysis can accurately quantify the error associated with the collection of samples which is governed by protocol, but remains varied by the discipline used as well as the individuals taking the samples. Furthermore, digital image processing and traditional sieving methods cannot be directly compared to one another because traditional methods are three-dimensional in nature and the digital images are two-dimensional. This is why a percent difference is used to compare the two rather than a percent error.

A sample plot is shown in figure 8 to give an example of how data points would look if error bars representing quantifiable error were included. If many data points were included in a single plot, it would become very crowded and trends would become difficult to discern; therefore, after this example, error bars will often be neglected in deference to clarity. The horizontal error bars represent the error propagated through the calculation of the mean grain size for the sieved grains. The most significant source of error that can be quantified is from the size of the grains retained on each sieve. A Gaussian distribution is assumed for each sieve size, which means each grain in that sieve is likely to be within 34% of the size that is exactly between that sieve and the one above it in the shaker. The vertical error bars are

generated by the algorithm in a similar fashion, and also assume a Gaussian distribution between the sieves. It should be noted that 200% error is not uncommon in this field of research, thus any improvements that bring quantifiable error to below 100% are significant.

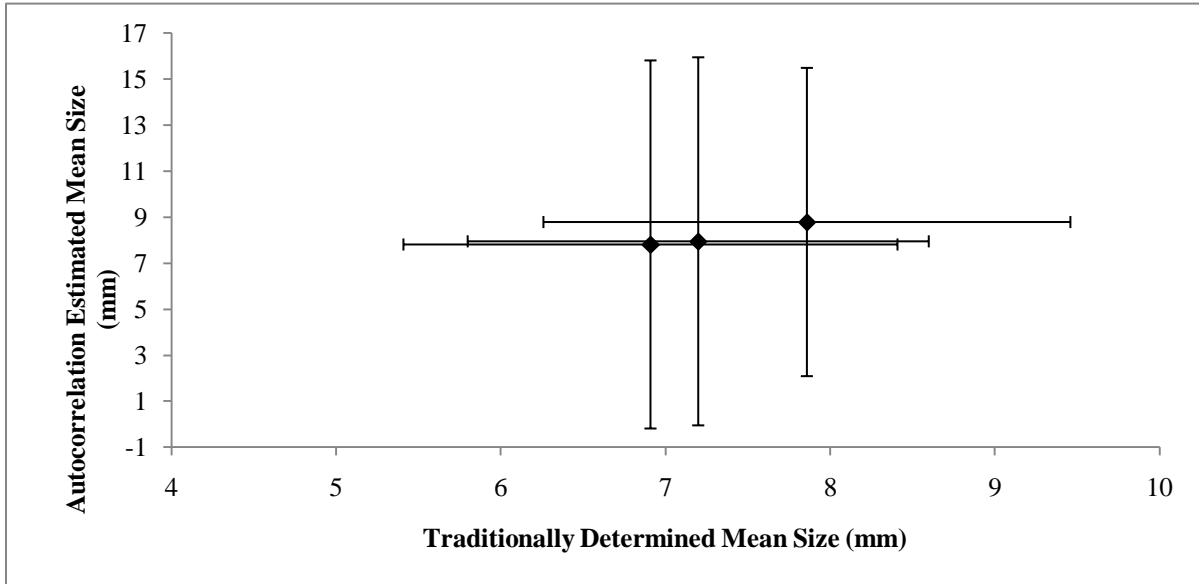


Figure 8. Plot of three data points with quantifiable error represented by error bars. Inclusion of these error bars in a plot with many data points would be distracting so they are left off of many of the graphs presented here in deference to clarity.

DATA AND RESULTS

Data

Data in the form of images were acquired from locations in the portion of Boone Creek that runs through Appalachian State University and from the New River at locations along the Greenway in Boone, NC. (Figure 1 in the Methods section shows a map of the approximate locations.) At each location in the New River, cross sectional sites were chosen to transverse the stream perpendicular to the banks. This provides grain size distribution information from one side of the stream to the other, thus enhancing the ability of a numerical model to properly characterize flow within the channel at a specific location. Though the physical separation between the sites is fairly large in this study (see figure 1) it would be a simple process to tighten the distances between sites and increase the number of sites to accommodate modeling preferences.

Table 4 is a partial list of the in situ data images collected from the New River each with corresponding information concerning focal length and grain size obtained from traditional sieving methods. A complete listing of all data images, from both the field and the lab, can be found in Appendix A. Some sample images of both calibration images and data images from the lab and the field can be viewed in Appendix B.

Table 4. A partial list of data images taken from the New River (summer 2008) showing the original camera image identifier, the focal length, the new identifier, and the sieved mean grain size of the sample taken at the time the image was acquired.

Camera Image Identification	Focal Length (equivalent mm)	Data Image Identification	Traditional Method Mean Grain Size
P6171105	200	NR616b01a2	4.87 mm
P6171112	200	NR616b01b2	7.86 mm
P6171118	200	NR616b01c2	14.76 mm
P6171125	200	NR616c01a2	10.02 mm
P6171134	200	NR616c01b2	14.84 mm
P6171153	200	NR616d01a2	4.51 mm
P6171168	200	NR616d01b2	17.30 mm
P6171180	200	NR616d01c2	20.95 mm
P6211241	200	NR620f01a2	24.79 mm
P6211252	200	NR620f01b2	43.85 mm
P6211263	200	NR620f01c2	16.30 mm
P6211273	200	NR620f01d2	9.59 mm
P6211284	200	NR620f01e2	3.94 mm
P6261289	200	NR625b01a2	1.97 mm
P6261294	200	NR625b01b2	9.82 mm
P6261299	200	NR625b01c2	15.24 mm
P6211188	200	NR620e01a2	19.55 mm
P6211198	200	NR620e01b2	25.92 mm
P6211209	200	NR620e01c2	6.91 mm
P6261322	200	NR625c01a2	9.08 mm
P6261327	200	NR625c01b2	24.73 mm
P6261337	200	NR625c01c2	24.66 mm
P6261344	200	NR625g01a2	23.27 mm
P6261360	200	NR625g01b2	23.15 mm

Optimization Results

To establish if more than one image of a particular site is required to accurately represent the stream bed surface, a set of five data images were taken of a single site. The camera remained centered over the same patch of sediment at 1 meter distance and 200mm magnification. The position of the frame was rotated slightly for each image. The images were then run, as single segments, through the autocorrelation algorithm. This process was repeated for eight other locations, all of which are included in the data set described above. For each set of five images per site, the autocorrelation algorithm provided nearly identical results with variances in the *tenths* or *hundredths* of a millimeter thus confirming that a single clear image of each site is sufficient to employ the autocorrelation routine for determining mean grain size.

Images obtained at the field locations showed significant patchiness – one portion of an image may be comprised of all small grains while another portion may be mostly larger grains. Figure 9 illustrates this behavior. To accommodate this patchiness in the automated algorithm, field images were segmented into 1, 4, 9, and 16 segments (if magnification allowed) and the results from each segment were averaged per image – these results were then compared to the results from the sieving method for each location. This was done for the full range of magnifications and maximum offset settings. When the percentage of automated estimations within tolerance is compared to the number of segments used for the data images, the accuracy as compared to sieving results, increases with segmentation up to 16 segments (figure 10). The values on the graph represent the average of the most accurate

estimations in the matrices (discussed in methods) for each of the four possible segmentations. The error bars represent the standard deviation of those values.



Figure 9. Variability in sediment distribution can significantly affect results, therefore segmenting the image and averaging the results for each segment provided improved accuracy overall.

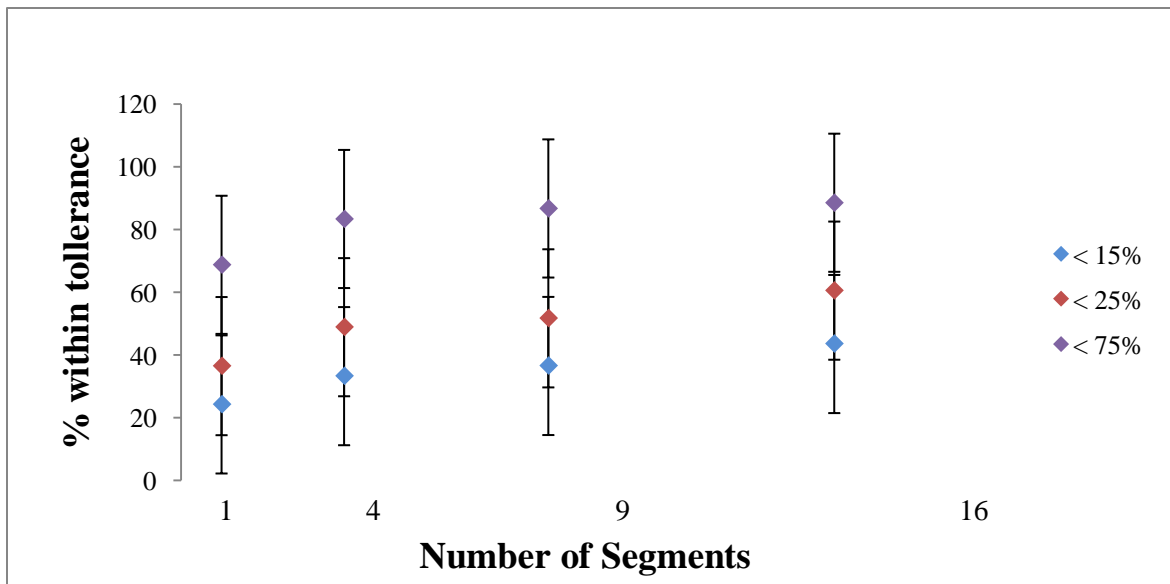


Figure 10. As the number of segments increases for data images, the accuracy of the method increases. This trend is more pronounced for the tighter tolerances.

As mentioned, segmenting the data images is desirable because the high variability in grain size inherent to high gradient streams is better represented by an average over several different sections of the image. One section of an image may differ considerably from another due to grouping of like sized grains caused by flow-induced segregation. Larger image areas would make this more important and may require a greater number of segments than is discussed here.

The results for calibration image segmentation are quite different, and conclusively show that the calibration images should not be heavily segmented. Figure 11 shows that as the number of segments for a calibration image increases, the accuracy decreases. (The same method as described for data segments was used to generate data points here.) A single segment calibration image generally outperformed the others. Only the highest resolution images showed both four and single segmented calibration images to be effective. It was

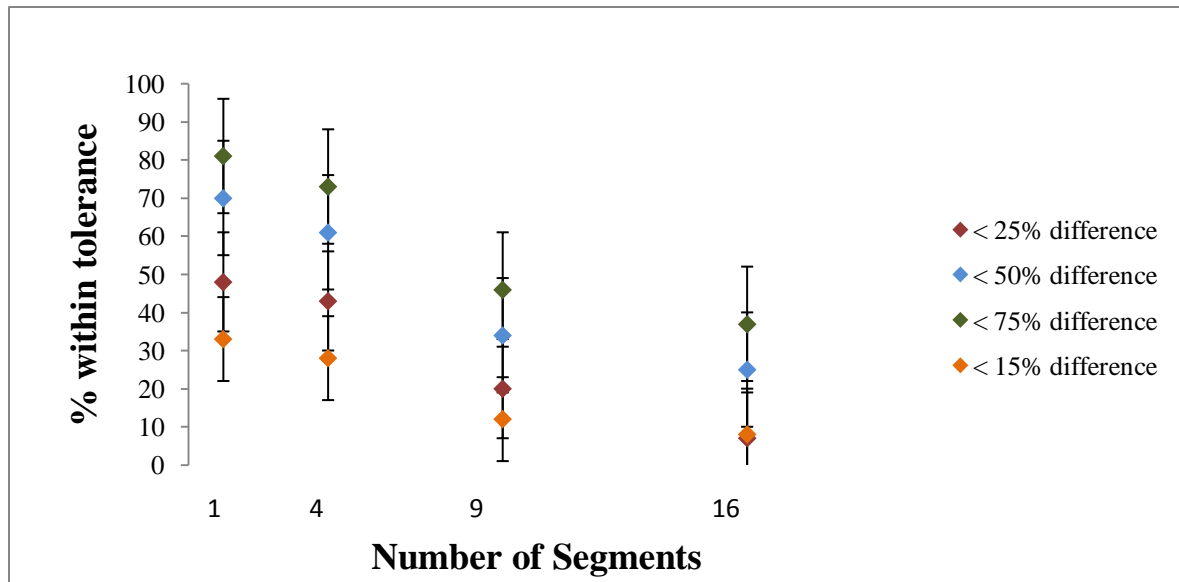


Figure 11. As the number of segments for a calibration image increases, accuracy decreases. Generally, it is recommended that the calibration image is used as a single segment.

noted that in *all* cases the number of segments used in the calibration routine needed to be the same or smaller than the number of segments used in the data images for optimal outcomes. As a general rule it is recommended that users employ single segment calibration images for all comparable applications. It is possible that segmenting the calibration images introduces unwanted noise to the correlation curves that subsequently impacts results.

The determination of mean grain size is independent of camera magnification for our set up of one meter distance from the streambed (figure 12). The magnifications were compared across all segmentations and maximum offsets. In figure 12, the best values chosen from each of the five matrices (discussed in methods) corresponding to a single magnification were averaged. Figure 12 only includes tolerances within 25% and 75% for

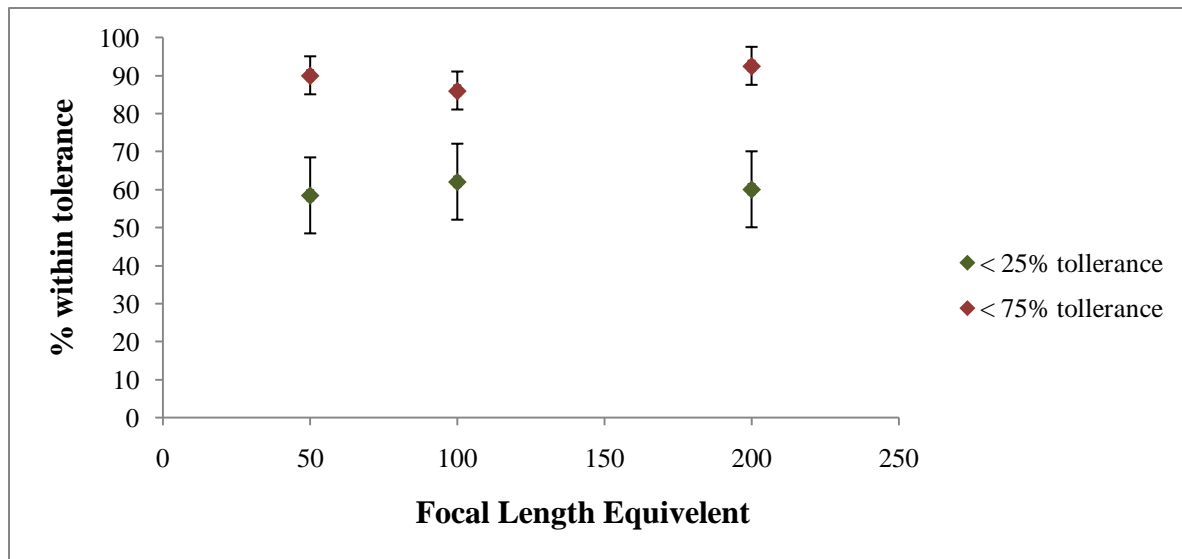


Figure 12. Results are reasonably consistent across all three focal lengths used. This allowed us to choose 200 mm as the preferred magnification which captures grain sizes down to 0.44 mm.

clarity. (All five tolerances can be viewed on a single plot in the appendix.) The error bars signify the standard deviation of the averaged values. This result allowed us to choose a magnification that best fit our needs. All of the data images reviewed in the final analysis

were taken at the 200mm focal length magnification to attain a minimum resolution of 0.44 mm grain size (at a minimum of 7 pixels per grain) which is within the suitable minimum size range for high gradient streams.

While the performance of the autocorrelation routine overall is independent of camera magnification, increased resolution does affect results for different mean grain sizes. When resolution is increased, the larger sizes are increasingly underestimated by the algorithm. This is an expected behavior due to smaller grains overlaying larger ones, thus ‘breaking up’ the larger grains to be interpreted as smaller grains by the autocorrelation routine (see figure 13). Conversely, lower resolution is in effect ‘blind’ to these small grains - at low magnification, small grains lose edge definition and are seen by the autocorrelation routine as blended into a single “large grain.” However, low resolution estimates the larger grains more



Figure 13. It is not unusual for some small grains to overlay larger ones in high gradient streams. This can cause the algorithm to underestimate the mean size of the image.

effectively. As a result, the choice of camera magnification for a given distance-to-target implies an optimal mean grain size and accompanying range for which the autocorrelation algorithm is most effective.

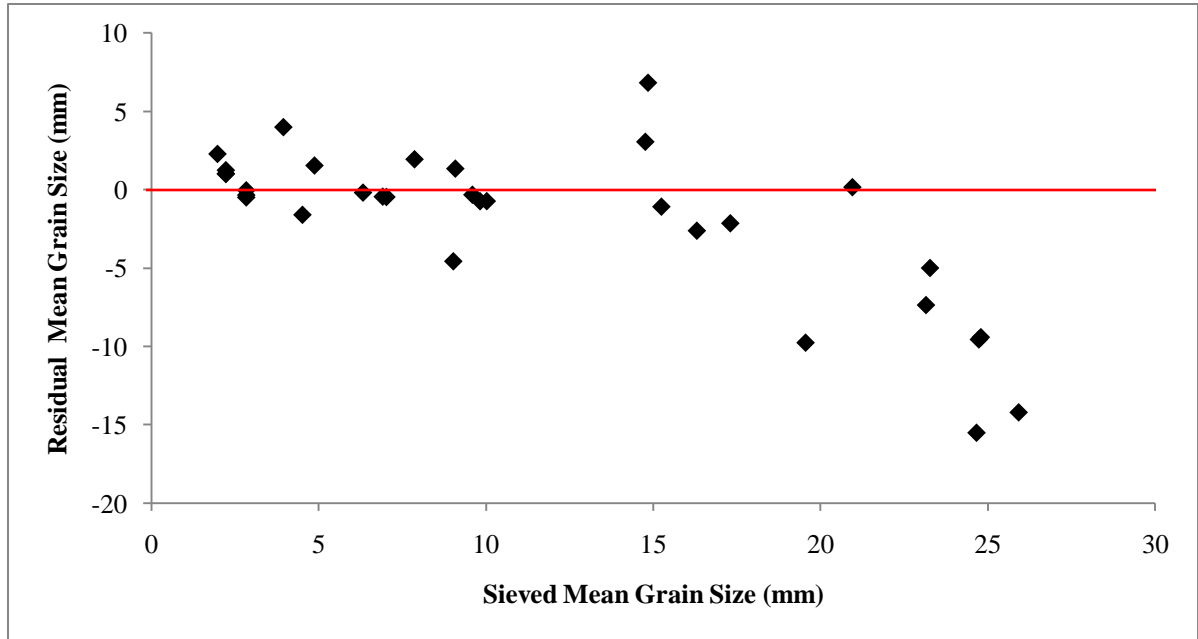


Figure 14. Plot of the residual (predicted – sieved) mean grain size. vs. sieved grain size [mm]. The algorithm tends to over predict the size of the small grains and under predict larger grains, implying an optimal grain size range.

When the autocorrelation algorithm estimated mean grain size is compared to traditional sieving results, this over- and under- representation of the grain size is confirmed. See figure 14, which shows the residual (autocorrelation-predicted – sieved) of mean grain size vs. sieved mean grain size for data images taken at 200 mm focal length, 100 maximum offset, and 16-segment data images. The range of “maximum offset” values used in the autocorrelation routine over which we can perform optimization analysis is linked to resolution and the degree of image segmentation. In order to properly represent all sediment in a sample mathematically, the maximum offset should technically include enough pixels to completely cross the largest grain in the sample. As stated earlier, we chose a minimum of 7

pixels to adequately represent a single grain, which corresponds to a minimum value for the “maximum offset” of 10. This is true regardless of magnification setting, although the higher the resolution, the greater the number of pixels per millimeter and thus the smaller the grains that can be resolved. Since the overall image size limits the degree of data image segmentation during processing, too many segments would limit the “image-width-to-process” computed as image width minus maximum pixel offset. Therefore, the maximum of the range of maximum offset values is limited to a smaller value for optimization scenarios with high degrees of image segmentation. In addition, if the maximum offset is too large the computational expense reduces the usefulness of the system. As a result, we limited the range of maximum offset from 10 to 800 pixels in our analysis.

The functional maximum value for the maximum offset ultimately limits the largest grains accessible for valid representation by the automated routine. The grain size range for high gradient streams is wide and can easily encompass grains from <1 mm to 100 mm in size or larger. For operational convenience in this study, the maximum grain size was initially limited to approximately 75 mm. However, to properly encompass such grains would require a 1200 pixel maximum offset, too large to be useful and outside our functional range as discussed above. As a result, results presented herein include samples and corresponding images in which the sieved mean grain size did not exceed 25 mm – although grains exceeding 800 pixels in width were included in some of these data, these were outliers whose representation, or misrepresentation, are well bounded by our error analysis.

Optimization of system performance based on maximum pixel offset show a possible range of useful maximum offsets for high gradient streams between 80 and 500 pixels. For

each offset considered, all useful data images were analyzed using the optimal data image segmentation of 16, calibration image segmentation of 1, and 200mm magnification. These images had a mean grain size of 2 mm to 25 mm. When the autocorrelation-estimated mean grain size was plotted against the traditionally determined mean grain size (discussed later), a linear best-fit line (using linear least squares) was generated with a projected slope and y-intercept. The slopes of the best fit lines for maximum pixel offsets ranging from 10 px to 500 px are shown in figure 15. A slope of one would indicate a perfect match to the one-to-one line where the estimated mean and the traditional mean were in agreement. Maximum pixel offsets of 80 or higher provided the best results. Note that maximum offsets greater than 500 were not accessible here due to image size limiting the size of the segments used. Larger image sizes would allow more in-depth investigation into this range; for example, at

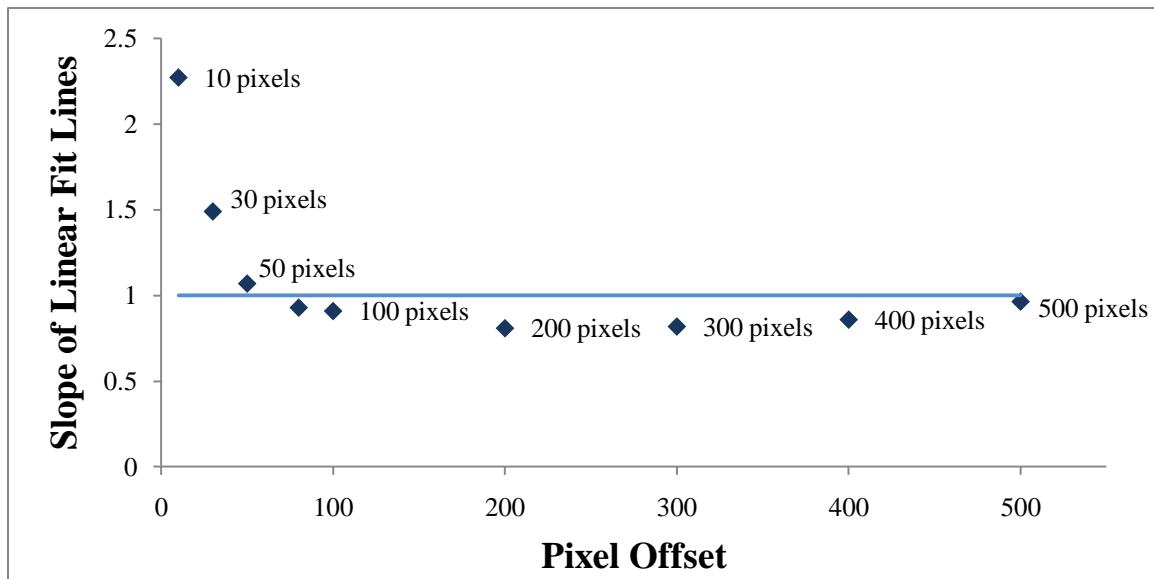


Figure 15. Plotting the slopes of different maximum pixel offsets against the one-to-one line illustrates the possibility of an optimum range of maximum offsets that give comparable results.

higher offsets, the linear fit function plotted in figure 15 may cross back above the x axis beyond 500 pixels.

It is possible that the maximum offset is not a critical factor once a minimum threshold has been achieved. It is interesting to note that this threshold seems to be far smaller than anticipated, close to 100 pixels as opposed to the several hundred that should technically be required to properly characterize larger grains in the sample. A 25 mm grain, for example, at a resolution of 16 pixels/mm should require approximately 400 pixels but in practice does not according to our results. One possible reason that the higher offsets did not perform better is that noise is introduced as the offset is increased. Noise may arise due to the larger spatial separation between the elements as well as differences in texture and color on large grains affecting the intensity values thus causing them to appear less correlated than they actually are.

An unexpected tolerance for high quality Joint Photographic Experts Group (JPEG, or JPG) compression was discovered during this research. In fact, when Raw (converted to TIF) images are used for the calibration files, and JPEG images are processed to determine mean grain size, the results are nearly identical to similar processing of raw data images at most maximum offsets. Originally, 29 images were compared using raw format and 88% compression JPEG format. Each image was processed at 19 different maximum offsets from 10 to 200 and 4 different image segmentations: 1, 4, 9, and 16. The only notable variances occurred when the data image was split into 16 segments and the maximum offset was set between 100 and 180 pixels. It was unclear why these particular offsets resulted in a measurable variance (up to 5 mm) while the others did not. These data were generated early in the development of the system and did not include higher maximum offsets.

An updated data set was generated to confirm JPEG compatibility and verify earlier results for higher maximum offsets. A total of 23 data images were run at 400 maximum offset and 16 image segments. The TIF file and JPEG file based results were within *hundredths* of a millimeter in *all* cases, thus reinforcing the evidence for invariance to JPEG compression (see figure 16). A complete table of the updated comparison values used to generate figure 16 can be found in the appendix.

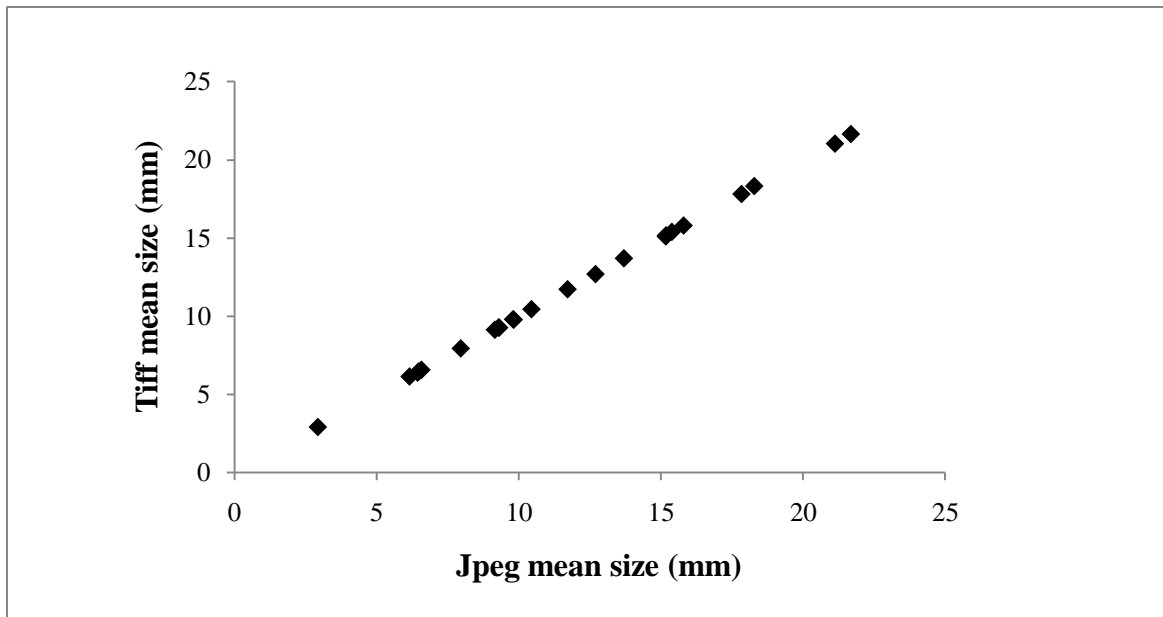


Figure 16. When data images processed in Tif format are compared to data images in Jpeg format the results are identical for all practical purposes. It was found that calibration images should always be in a lossless format.

One possible explanation for this compatibility may be that as a rule JPEG compression preferentially discards color information over light intensity information in an image. Our algorithm, which converts all images to grayscale, depends solely on the light intensity information. This unforeseen development enhances the usefulness of the system for those who wish to use the JPEG format either for its memory-saving benefits or simply for speed and cost effectiveness. Analysis did demonstrate that the use of a lossless image format for the calibration files is required.

The invariance of the autocorrelation algorithm to JPEG compression allowed for the inclusion of ten additional images to our field data set, all of which were taken in the JPEG format. These images were all from Boone Creek, and were taken in the summer of 2007. A complete listing of all images can be found in the tables of Appendix A.

Results Under Optimized Conditions

As mentioned earlier, results do indicate the possibility of an optimal grain size range for this system which is between 2 mm and 25 mm. As a rule, the size of grains less than 4 mm as determined by sieving was over-predicted by the algorithm, while the size of grains greater than 25 mm as determined by sieving was under-predicted by the algorithm *all* of the time, with a mix of over and under prediction for grain sizes between 4mm and 25 mm as determined by sieving. Figure 14 shows this tendency to under or over estimate the mean grain size. In general, in the size range between 4 and 25 millimeters, the algorithm underestimated the size 39% of the time and overestimated the mean size 61% of the time.

The methodology employed to acquire data in wetted conditions generated comparable results to dry data in the lab. The mean grain size predictions produced by the algorithm from the lab images, at a tolerance of 50% difference from mean grain sizes between 2 and 20 mm obtained from traditional methods, were accurate 90% of the time. The mean grain size predictions produced by the algorithm from the field data images, at a tolerance of 50% difference from mean grain sizes between 2 and 25 mm obtained from traditional methods, were accurate 82% of the time. Table 5 illustrates all the tolerances considered by this study for both the lab images and the field data images. Though a more robust data set would strengthen this comparability (both of these data sets were fairly small, 20 lab images and 38 field data images for a total of 58 images), our findings suggest that images obtained in the field perform comparably well for mean grain size predication as

images of the same sediment that has been removed, dried, and re-photographed in lab conditions. A larger data set would strengthen this conclusion.

Table 5. Comparison of Lab acquired data to field acquired data for each tolerance level considered.

TOLERANCE:	25%	50%	75%	100%
LAB	60%	90%	90%	90%
FIELD	50%	82%	97%	97%

When all 58 of the images that fall within the preferred grain size range are included in the data set and optimized parameters used, the mean grain size can be determined within 50% of the traditionally determined mean 85% of the time using the autocorrelation method. If a 75% tolerance is acceptable, this percentage increases to 97% accuracy. These values were arrived at using the optimal settings as determined by earlier analysis. Three maximum offsets were chosen to illustrate the range discussed earlier: 100, 300, and 500 maximum offset. The remaining parameters used are: 1 segment calibration images, 16 segment data images, and 200mm camera magnification which translates to 16 pixels per millimeter resolution.

Figure 17 is a plot of the mean grain size predicted by the autocorrelation routine vs. the mean grain size obtained from traditional sieving methods for all 58 of the images that fall within the preferred grain size range. The 1:1 line representing a perfect match between autocorrelation and sieving methods is included for reference. Also included are 68% confidence interval lines (one standard deviation) for the data about the 1:1 line. In general, comparison of the two methods shows that the two methods differ by approximately 200% -

sometimes much better, sometimes worse. However, due to the high error associated with standard sieving methods, these results do not make claims to the ability or inability of the autocorrelation method to accurately represent the actual two-dimensional surface of grains, and may actually outperform standard methods for surface grain size characterization.

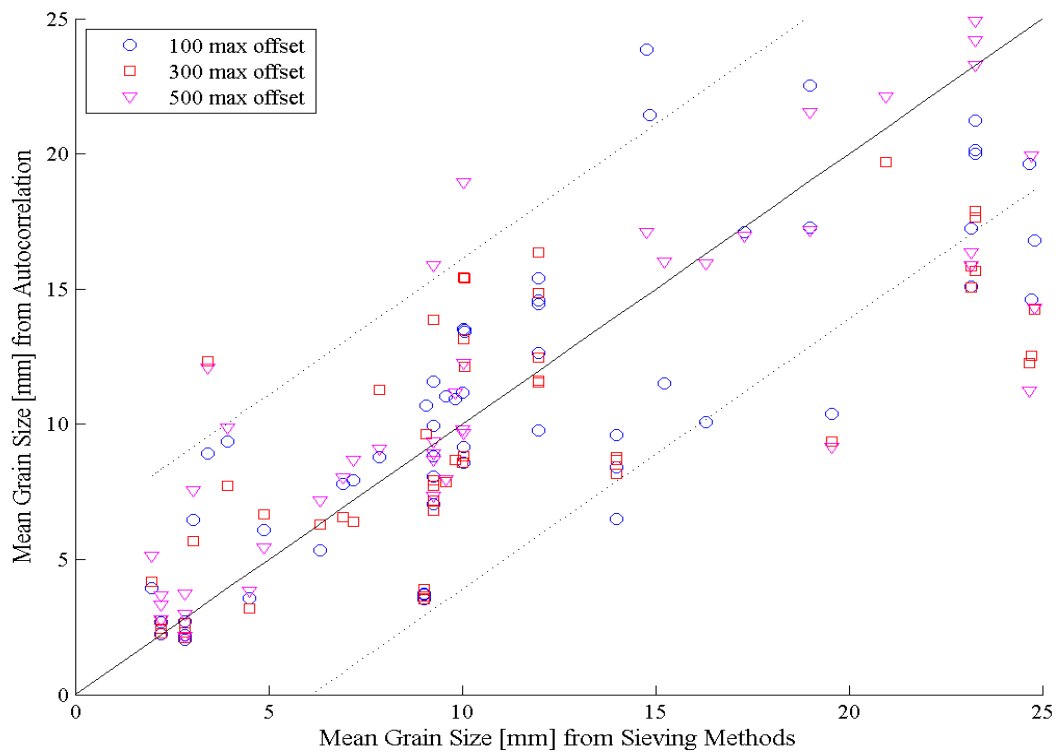


Figure 17. The mean grain size predicted by the autocorrelation routine vs. the mean grain size obtained from traditional sieving methods for all 58 of the images that fall within the preferred grain size range. The 1:1 line (solid) representing a perfect match between autocorrelation and sieving methods is included for reference. Also included are 68% confidence interval lines (dashed- one standard deviation) for the data about the 1:1 line.

Results for the mean grain size at the locations along Boone Creek and the New River listed on the map presented earlier are now included on that map in figure 18. The results for The New River are the average of the cross-sectional locations at each site. Boone Creek results are from a single location at each site listed. One location, along Boone Creek, did not have useful results due to out of focus images. These predictions are well within the expected grain size range for the New River and Boone Creek.

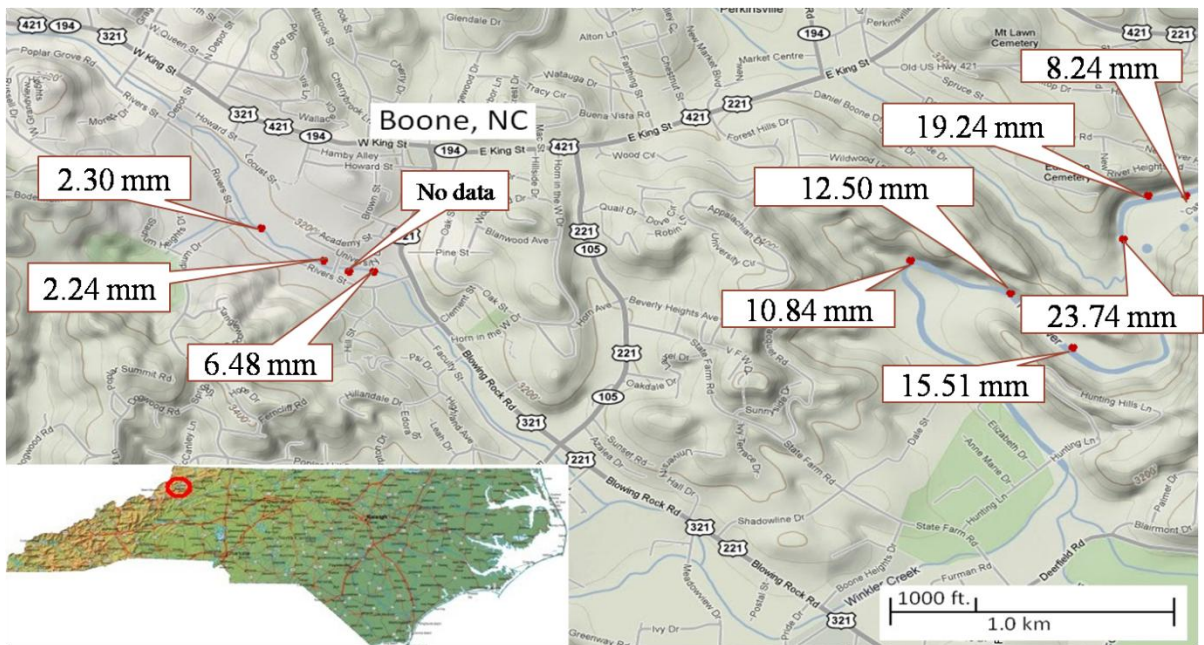


Figure 18. Average mean grain size for locations specified earlier. Boone Creek results are from single locations at each site. New River results are the averaged mean size values for the cross-sectional locations at each site.

CONCLUSIONS AND RECOMMENDATIONS

The autocorrelation grain sizing method discussed herein can be used for determining the mean grain size in high gradient streams and rivers with skill if the optimized parameters for calibration, segmentation, maximum offset, and magnification are employed in conjunction with an order-of-magnitude estimation of the expected mean grain size. Within a tolerance of 50% difference from traditional methods, the mean grain size is accurately determined 85% of the time using the proposed method. This was achieved using a single segment calibration image, sixteen segment data images, a maximum offset of 100 pixels, a magnification of 200mm focal length at a distance from target of 1m for data images of sediment with a mean size between 2 mm and 25 mm. The speed with which data may be collected and processed makes application of this method preferable for some purposes due to the potential for a dramatic increase in temporal resolution.

The lower end of the ideal mean grain size range (between 2 mm and 25 mm) is constrained by the available camera magnification (resolution): if a grain is the same size or smaller than a few pixels the algorithm cannot effectively resolve it as a separate grain. The upper end of the ideal mean grain size range is defined in part by the overall image size and the degree of image segmentation for processing and by the three-dimensional layering of the stream bed: Frequently, smaller grains overlie larger ones decreasing the mean grain size predicted by the autocorrelation routine as compared to the actual mean. However, generation of a more sizeable data set for mean grain sizes larger than 25 millimeters would

be required to confirm or expand this range. General overestimation of grain size within the preferred range needs to be investigated more thoroughly as well. It is unclear whether or not these estimation errors are a result of the spatial autocorrelation or the non-negative least squares fitting routine used to determine the mean size, or an actual under prediction of surface grain size distribution from standard sampling and sieving methods

It is conceivable that the presented autocorrelation routine is more accurate than our results imply due to the fact we are comparing a system that uses two dimensional digital data to a method that uses three dimensional sieved data. The demonstrated skill of the proposed method strongly suggests that it may be quite accurate for the two dimensional images and thus for the surface of the streambed. The only way to properly assess this would be to physically count and measure every grain in many different images and compare the results to those of the system. This is an option outside the scope of this investigation but certainly a possibility for future consideration.

A greater understanding of how calibration affects the results is necessary for expansion of use. It may be possible to define an overall calibration set that would encompass a variety of uses, or perhaps to define unique calibration sets that remain broad in their application. Investigation into the optimization of the number of calibration curves – the span and dimensionality of the basis set of images - would be a useful exercise for this system as well.

It is also not apparent what effect color variation has on results; therefore it is recommended that users calibrate the system with sediment similarly to that which is to be tested. Color variations in the sediment could cause the autocorrelation routine to

inadvertently lump more than one grain with another of identical color, particularly if the grains are of similar grayscale intensity and are close to the minimum size range allowed by resolution. Variations of color on a single large aggregate grain could cause it to appear as two or more grains, thus resulting in an underestimation of the true size for purposes of hydrodynamic coupling. For this study, we assumed the errors caused by this to be randomized and accounted for by calibrating the system with sediment from the same body of water from which we acquired data.

We did not specifically explore the effect of light intensity in the image on results, but suspect there is likely to be some effect on the system, particularly if the light intensity is very high throughout the image. A flash was not used on in situ data due to reflection issues with the Plexiglas box, therefore one was not used for the lab data images acquired for parameter optimization either. This meant all the images had a relatively low light intensity value overall, particularly the in situ images taken early in the day or on cloudy days. However, for thoroughness, all the lab images were taken both with and without a flash - preliminary results using the data images taken with the flash did show a measureable difference from those taken without a flash. The effects of light intensity is an issue requiring further scrutiny and may ultimately give greater insight into the limitations of the method presented.

It is important to note that use of this system does require a certain attention to detail and a comprehensive grasp of how it functions when collecting and processing data. One must be adept at acquiring and preparing useful images. This includes avoiding images of sediment covered with interfering biological material such as mud, moss, leaves, fish etc. It

is also wise to have a reasonable approximation of what the mean grain size should be in a given stream or cross section so that images giving unusual results can be examined more closely for anomalies. Alert users will be able to see which images may be routinely over or under represented by the method and could possibly make adjustments to modeling parameterizations accordingly.

More detailed investigation into the optimum maximum offset and its effect on accuracy is also recommended. The fact that the system begins to return good results at maximum offsets that are far smaller than those required to cross the largest grains in the sample was an unexpected result without an immediate explanation. Lower offsets are processed far more quickly than higher ones, so are potentially a benefit for the system user who requires a quick turnaround of results. Due to resolution issues, it was impossible to thoroughly investigate higher offsets at the optimal data image segmentation. However, understanding this behavior may lead to identification of a clear, reproducible mathematical relationship between the mean grain size, optimal resolution, and maximum pixel offset. Quantification of that relationship would make this system scalable for practically any purpose, including planetary surface characterization, for example.

A different method for eigenvalue decomposition, other than non-negative least squares, may yield more accurate results at higher offsets. It is also possible that the system may perform optimally at offsets much higher than those tested here, something that larger image sizes could confirm or deny. It is equally possible that none of these would affect the results at higher offsets and that a threshold value will always remain true. More investigation is needed to clarify this issue.

Suggested improvements to the apparatus include attaching the Plexiglas ‘box’ to the frame so that the user need not stabilize the box relative to the frame while obtaining images. Also, at the risk of higher cost, a camera with considerably higher resolution could be used to increase the number of pixels/mm, thus expanding its potential grain size range. Crafting a method for moving the camera closer to the streambeds’ surface could achieve similar resolution improvement.

While this study has effectively demonstrated the potential usefulness of employing autocorrelation techniques on digital images of sediment from high gradient streams to determine mean grain size, there remain a considerable number of questions to be answered. It is the author’s hope that future research will better define both the strengths and limitations of this method toward a reliable, inexpensive, and rapid methodology to determine mean size for a wide variety of applications.

WORKS CITED

- ASTM D421-85, 2007, Standard Practice for Dry Preparation of Soil Samples for Particle Size Analysis and Determination of Soil Constants: American Society for Testing and Materials, 2007.
- Adams, J., 1979, Gravel size analysis from photographs: American Society of Civil Engineers: Proceedings, Journal of the Hydraulics Division, v. 104, p. 1247–1255.
- Buscombe, D., 2008, Estimation of grain-size distributions and associated parameters from digital images of sediment: Sedimentary Geology, v. 210, p. 1-10.
- Buscombe, D., and Gerhard M., 2009, Grain-size information from the statistical properties of digital images of sediment: Sedimentology, v. 56, p. 421-438.
- Butler, J. B., and Stuart N. L., and Chandler, J.H., 2001, Automated extraction of grain-size data from gravel surfaces using digital image processing: Journal of Hydraulic Research v. 39.5, p. 519-529.
- ERDAS, Inc., 1999: ERDAS Field Guide, 5th Edition: Atlanta, GA, ERDAS, Inc. 204-208 p.
- Fishbach A., Nelken I., Yeshurun Y., 2001, Auditory edge detection: A neural model for physiological and psychoacoustical responses to amplitude transients: Journal of Neurophysiology, v. 85(6), p. 2303-2323.
- Fredsoe, J. and Deigaard, R., 1992, Mechanics of Coastal Sediment Transport, Advanced Series on Ocean Engineering Vol. 3: Singapore, World Scientific, 194-226 p.
- Gasparini, N.M., Tucker, G.E., Bras, R.L., 1999, Downstream fining through selective particle sorting in an equilibrium drainage network, Geology, v. 27(12), p. 1079-1082.
- Gluschke, M., Voitke, P., Wellnitz, J., and Lepom, P., 2004, Sieving of sediments for subsequent analysis of metal pollution: results of a German interlaboratory study: Accred Qual Assur 2004, v. 9, p. 624–628.
- Haan, C.T., Barfield B.J. and Hayes, J.C., 1994, Design Hydrology and Sedimentology for Small Catchments: San Diego, CA, Academic Press Inc, 204-236 p.

- Ibbeken, H. and Schleyer, R., 1984, Photo-sieving: a method for grain-size analysis of coarse grained, unconsolidated bedding surfaces: *Earth Surface Processes and Landforms*, v. 11, p.59–77.
- Loveland, P.J., and Whalley, W.R., 2001, *Soil and Environmental Analysis: Physics Methods*: New York, Marcel-Dekker, 281-296 p.
- McEwan, I.K., Sheen T.M., Cunningham, G.J, and Allen A.R., 2000, Estimating the size composition of sediment surfaces through image analysis: *Proceedings of the Institution of Civil Engineers – Water, Maritime, and Energy*, v. 142(4), p. 189-195.
- Meer P. and Georgescu B., 2001, Edge detection with embedded confidence: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 23(12), p. 1351-1365.
- Nielsen, P., 1992, *Coastal Bottom Boundary Layers and Sediment Transport*, Advanced Series on Ocean Engineering, Vol. 4: Singapore, World Scientific, 95-128 p.
- Nino, Y., Lopez, F., and Garcia, M., 2003, Threshold for particle entrainment into suspension: *Sedimentology*, v. 50, p. 247–263.
- Penders, C.A. and Thaxton, C.S., 2010 - In preparation, Surface Characterization using Autocorrelation Techniques on Digital Images: To be submitted to the *Journal of Sedimentary Research*, SEPM.
- Rubin, D. M., 2004, A simple autocorrelation algorithm for determining grain size from digital images of sediment: *Journal of Sedimentary Research*, v. 74(1), p. 160-165.
- Rubin D.M., Chezar H., Harney J.N., Topping D.J., Melis T.S., and Sherwood C.R., 2007, Underwater Microscope for measuring spatial and temporal changes in bed-sediment grain size: *Sedimentary Geology*, v. 202, p. 402-408.
- Sime, L.C. and Ferguson, R.I., 2003, Information on Grain Sizes in Gravel-Bed Rivers by Automated Image Analysis: *Journal of Sedimentary Research*, v. 73(4), p. 630-636.
- USDA 1979, *Field manual for research in agricultural hydrology*, USDA Agriculture Handbook No. 224: Washington, DC, U.S., Government Printing Office.
- van Dokkum P.G., 2001, Cosmic-ray rejection by Laplacian edge detection: *Publications of the Astronomical Society of the Pacific*, v. 113(789), p. 1420-1427.
- Woodman R.J., Playford D.A., Watts G.F., Cheetham C., Reed C., Taylor R.R., Puddey I.B., Beilin L.J., Burke V., Mori T.A., Green D., 2001 Improved analysis of brachial artery ultrasound using a novel edge-detection software system: *Journal of Applied Physiology*, v. 91(2), p. 929-937.

Zhou, Y., Starkey, J., and Mansinha, L., 2004, Segmentation of petrographic images by integrating edge detection and region growing: *Computers & Geosciences*, v. 30(8), p. 817-831.

Appendix A

Data Tables

Table A- 1. Image dimensions for the Olympus SP-510UZ Camera

Magnification	Width (cm)	Length (cm)	Area (cm ²)
38 mm	90.3	66.9	6029
50 mm	70	52.3	3658
100 mm	31.7	25.3	802
200 mm	18.7	13.3	248
3890 mm	11.9	8.9	106

Table A- 2. Estimated mean grain size values for Tiff and Jpeg image files with the difference between them.

Tiff estimated mean grain size (mm)	Jpeg estimated mean grain size (mm)	Difference in estimated mean size (mm)
6.42635843	6.397855	-0.0285
9.81432136	9.77978	-0.03454
17.8314381	17.81333	-0.01811
9.30045713	9.273127	-0.02733
21.68155636	21.65048	-0.03108
2.91788892	2.917197	-0.00069
15.16560176	15.10435	-0.06125
21.12215415	21.03527	-0.08688
15.3802888	15.39692	0.01663
12.69229968	12.6941	0.001803
13.69161699	13.70338	0.011764
9.28015079	9.279315	-0.00084
7.94819106	7.943625	-0.00457
9.79003664	9.8046	0.014563

11.70956601	11.726	0.016434
6.47756372	6.4799	0.002336
6.14487826	6.1472	0.002322
6.56622266	6.5743	0.008077
10.43674002	10.45014	0.013401
15.17698077	15.17253	-0.00445
9.147936	9.134593	-0.01334
18.28037758	18.32188	0.041507
15.79358395	15.79928	0.0057

Table A- 3. A list of all images acquired in the lab for optimization purposes, including the sieved mean grain size.

Cameral Image Identification	Focal Length (equivalent mm)	Data Image Identification	Sieved Mean Grain Size
P7101970	50	LB01_50mm	3.04 mm
P7101972	100	LB01100mm	3.04 mm
P7101975	200	LB01200mm	3.04 mm
P7101978	50	LB02_50mm	3.44 mm
P7101980	100	LB02100mm	3.44 mm
P7101982	200	LB02200mm	3.44 mm
P7101986	50	LB03_50mm	10.60 mm
P7101988	100	LB03100mm	10.60 mm
P7101990	200	LB03200mm	10.60 mm
P7101994	50	LB04_50mm	10.60 mm
P7101996	100	LB04100mm	10.60 mm
P7101998	200	LB04200mm	10.60 mm

P7102002	50	LB05_50mm	9.26 mm
P7102004	100	LB05100mm	9.26 mm
P7102006	200	LB05200mm	9.26 mm
P7102010	50	LB06_50mm	9.26 mm
P7102012	100	LB06100mm	9.26 mm
P7102014	200	LB06200mm	9.26 mm
P7152019	50	LB07_50mm	9.26 mm
P7152021	100	LB07100mm	9.26 mm
P7152023	200	LB07200mm	9.26 mm
P7152025	50	LB08_50mm	9.26 mm
P7152027	100	LB08100mm	9.26 mm
P7152029	200	LB08200mm	9.26 mm
P7152031	50	LB09_50mm	9.26 mm
P7152033	100	LB09100mm	9.26 mm
P7152035	200	LB09200mm	9.26 mm
P7152037	50	LB10_50mm	10.04 mm
P7152039	100	LB10100mm	10.04 mm
P7152041	200	LB10200mm	10.04 mm
P7152043	50	LB11_50mm	10.04 mm
P7152045	100	LB11100mm	10.04 mm
P7152047	200	LB11200mm	10.04 mm
P7152049	50	LB12_50mm	10.04 mm
P7152051	100	LB12100mm	10.04 mm
P7152053	200	LB12200mm	10.04 mm
P7152055	50	LB13_50mm	11.98 mm
P7152057	100	LB13100mm	11.98 mm
P7152059	200	LB13200mm	11.98 mm
P7152061	50	LB14_50mm	11.98 mm
P7152063	100	LB14100mm	11.98 mm

P7152065	200	LB14200mm	11.98 mm
P7152067	50	LB15_50mm	11.98 mm
P7152069	100	LB15100mm	11.98 mm
P7152071	200	LB15200mm	11.98 mm
P7152073	50	LB16_50mm	11.98 mm
P7152075	100	LB16100mm	11.98 mm
P7152077	200	LB16200mm	11.98 mm
P7152079	50	LB17_50mm	11.98 mm
P7152081	100	LB17100mm	11.98 mm
P7152083	200	LB17200mm	11.98 mm
P7152085	50	LB18_50mm	19.00 mm
P7152087	100	LB18100mm	19.00 mm
P7152089	200	LB18200mm	19.00 mm
P7152091	50	LB19_50mm	19.00 mm
P7152092	100	LB19100mm	19.00 mm
P7152095	200	LB19200mm	19.00 mm
P7152097	50	LB20_50mm	19.00 mm
P7152099	100	LB20100mm	19.00 mm
P7152101	200	LB20200mm	19.00 mm
P7152103	50	LB21_50mm	19.00 mm
P7152105	100	LB21100mm	19.00 mm

All of the images listed in the following table were taken from the New River.

Table A- 4. A list of all images considered from the New River. The sieved mean grain size is from the sample collected from the image area at the time the image was taken.

Camera Image Identification	Focal Length (equivalent mm)	Data Image Identification	Traditional Method Mean Grain Size
P6171105	200	NR616b01a2	4.87 mm
P6171112	200	NR616b01b2	7.86 mm
P6171118	200	NR616b01c2	14.76 mm
P6171125	200	NR616c01a2	10.02 mm
P6171134	200	NR616c01b2	14.84 mm
P6171153	200	NR616d01a2	4.51 mm
P6171168	200	NR616d01b2	17.30 mm
P6171180	200	NR616d01c2	20.95 mm
P6211241	200	NR620f01a2	24.79 mm
P6211252	200	NR620f01b2	43.85 mm
P6211263	200	NR620f01c2	16.30 mm
P6211273	200	NR620f01d2	9.59 mm
P6211284	200	NR620f01e2	3.94 mm
P6261289	200	NR625b01a2	1.97 mm
P6261294	200	NR625b01b2	9.82 mm
P6261299	200	NR625b01c2	15.24 mm
P6211188	200	NR620e01a2	19.55 mm
P6211198	200	NR620e01b2	25.92 mm
P6211209	200	NR620e01c2	6.91 mm
P6211220	200	NR620e01d2	6.32 mm
P6211230	200	NR620e01e2	7.02 mm
P6261322	200	NR625c01a2	9.08 mm
P6261327	200	NR625c01b2	24.73 mm
P6261337	200	NR625c01c2	24.66 mm

P6261344	200	NR625g01a2	23.27 mm
P6261360	200	NR625g01b2	23.15 mm

Table A- 5. A list of all images from Boone Creek included in this study. All but the final three in the list were taken in the Jpeg file format.

Camera Image Identification	Focal Length (equivalent mm)	Data Image Identification	Traditional Method Mean Grain Size
P5250803	200	BC524b01a2	2.22 mm
P5250818	200	BC524b16a2	2.22 mm
P5250822	200	BC524b20a2	2.22 mm
P5250835	200	BC524c11a2	9.02 mm
P5250842	200	BC524c18a2	9.02 mm
P5250850	200	BC524c26a2	9.02 mm
P5250862	200	BC524d08a2	2.83 mm
P5250867	200	BC524d13a2	2.83 mm
P5250879	200	BC524d25a2	2.83 mm
P6070902	200	BC606f01a2	13.99 mm
P6070903	200	BC606f02a2	13.99 mm
P6070904	200	BC606f03a2	13.99 mm

Table A-6. 50 mm matrix values for 25% difference from sieved mean grain size. From right to left, top to bottom the maximum offsets are: 200, 300, 400, 500, and 600 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	33	43	67	57
4	0	29	38	33
9	0	9	5	9
16	0	0	0	5

Data Image Segments				
Calibration Segments:	1	4	9	16
1	33	43	67	57
4	0	33	43	52
9	0	0	0	5
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	29	53	57	x
4	0	19	x	x
9	x	x	x	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	29	52	x	x
4	x	x	x	x
9	x	x	x	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	29	57	x	x
4	x	x	x	x
9	x	x	x	x
16	x	x	x	x

Table A-7. 100 mm matrix values for 25% difference from sieved mean grain size. From left to right, top to bottom the maximum offsets are 300, 400, 500, 600, and 700 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	40	45	60	70
4	55	55	55	65
9	40	40	40	35
16	30	35	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	35	45	60	70
4	55	60	60	65
9	20	25	40	30
16	10	5	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	35	45	55	65
4	45	45	60	45
9	15	20	30	20
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	35	50	55	x
4	40	45	40	x
9	0	5	15	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	45	45	50	x
4	20	25	x	x
9	x	x	x	x
16	x	x	x	x

Table A-8. 200 mm matrix values for 25% difference from sieved mean grain size. From left to right, top to bottom the maximum offsets are: 400, 500, 600, 700, and 800 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	45	60	40	45
4	55	70	65	65
9	30	40	30	0
16	15	10	5	0

Data Image Segments				
Calibration Segments:	1	4	9	16
1	50	55	45	60
4	45	55	50	50
9	25	20	15	30
16	5	20	5	30

Data Image Segments				
Calibration Segments:	1	4	9	16
1	45	50	40	x
4	50	45	25	x
9	15	10	5	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	35	45	25	x
4	35	50	30	x
9	15	20	15	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	30	50	x	x
4	40	45	x	x
9	x	x	x	x
16	x	x	x	x

Table A-9. 50 mm matrix values for 50% difference from sieved mean grain size. From left to right, top to bottom the maximum offsets are: 200, 300, 400, 500, and 600 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	43	67	76	86
4	19	38	67	62
9	0	24	33	19
16	0	10	0	5

Data Image Segments				
Calibration Segments:	1	4	9	16
1	48	71	76	86
4	52	52	71	76
9	0	10	14	14
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	57	71	76	x
4	5	48	x	x
9	x	x	x	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	57	76	x	x
4	x	x	x	x
9	x	x	x	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	67	90	x	x
4	x	x	x	x
9	x	x	x	x
16	x	x	x	x

Table A-10. 100 mm matrix values for 50% difference from sieved mean grain size. From left to right, top to bottom the maximum offsets are: 300, 400, 500, 600, and 700 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	60	60	75	85
4	60	65	65	80
9	50	50	55	65
16	45	40	40	40

Data Image Segments				
Calibration Segments:	1	4	9	16
1	55	75	70	85
4	60	75	65	80
9	45	45	55	55
16	25	30	35	30

Data Image Segments				
Calibration Segments:	1	4	9	16
1	55	65	75	85
4	55	70	60	75
9	40	40	55	50
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	55	65	65	x
4	55	65	65	x
9	15	20	30	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	55	65	65	x
4	40	55	40	x
9	x	x	x	x
16	x	x	x	x

Table A-11. 200 mm matrix values for 50% difference from sieved mean grain size. From left to right, top to bottom the maximum offsets are: 400, 500, 600, 700, and 800 pixels.

Data Image Segments				
Calibration Segments:	1	4	9	16
1	75	80	80	70
4	80	85	80	80
9	45	60	65	15
16	15	15	15	15

Data Image Segments				
Calibration Segments:	1	4	9	16
1	80	75	75	70
4	75	70	70	70
9	25	30	40	70
16	15	35	15	70

Data Image Segments				
Calibration Segments:	1	4	9	16
1	75	80	60	x
4	65	65	55	x
9	20	15	15	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	75	75	60	x
4	55	75	45	x
9	15	35	25	x
16	x	x	x	x

Data Image Segments				
Calibration Segments:	1	4	9	16
1	65	70	x	x
4	50	65	x	x
9	x	x	x	x
16	x	x	x	x

Appendix B

Images

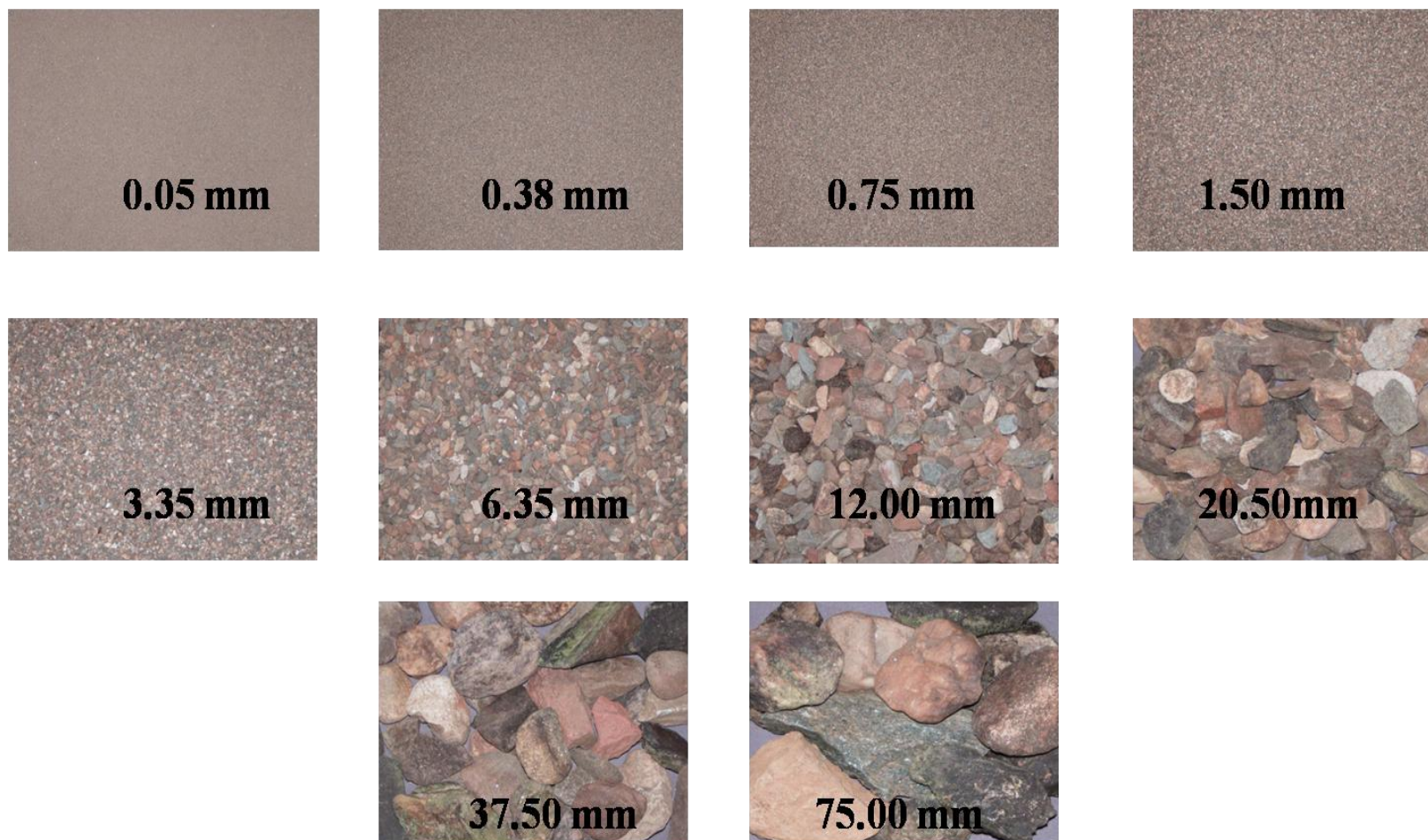


Figure 1. Calibration images used to generate calibration curves used by the autocorrelation algorithm. These are for 50 mm focal length.

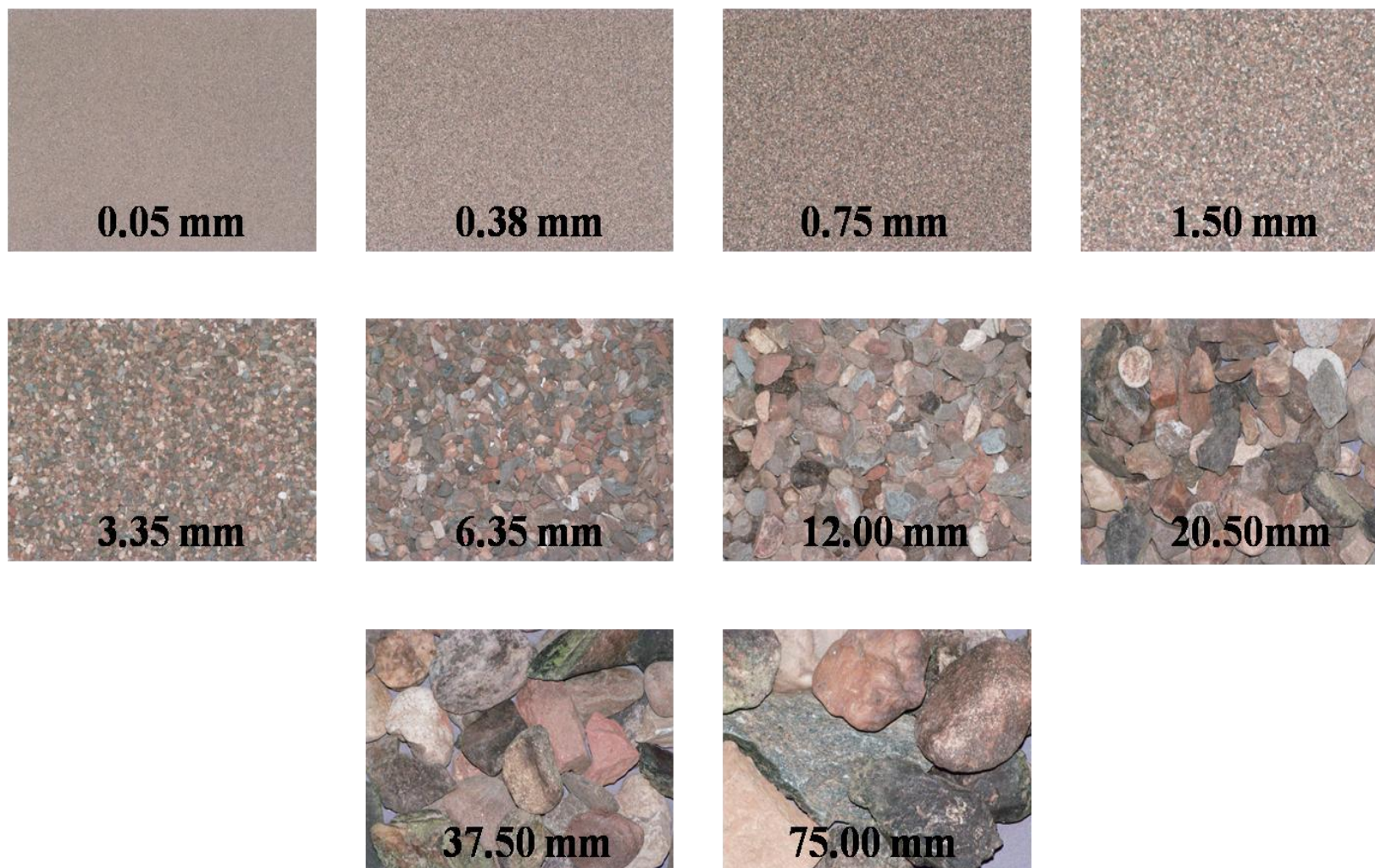
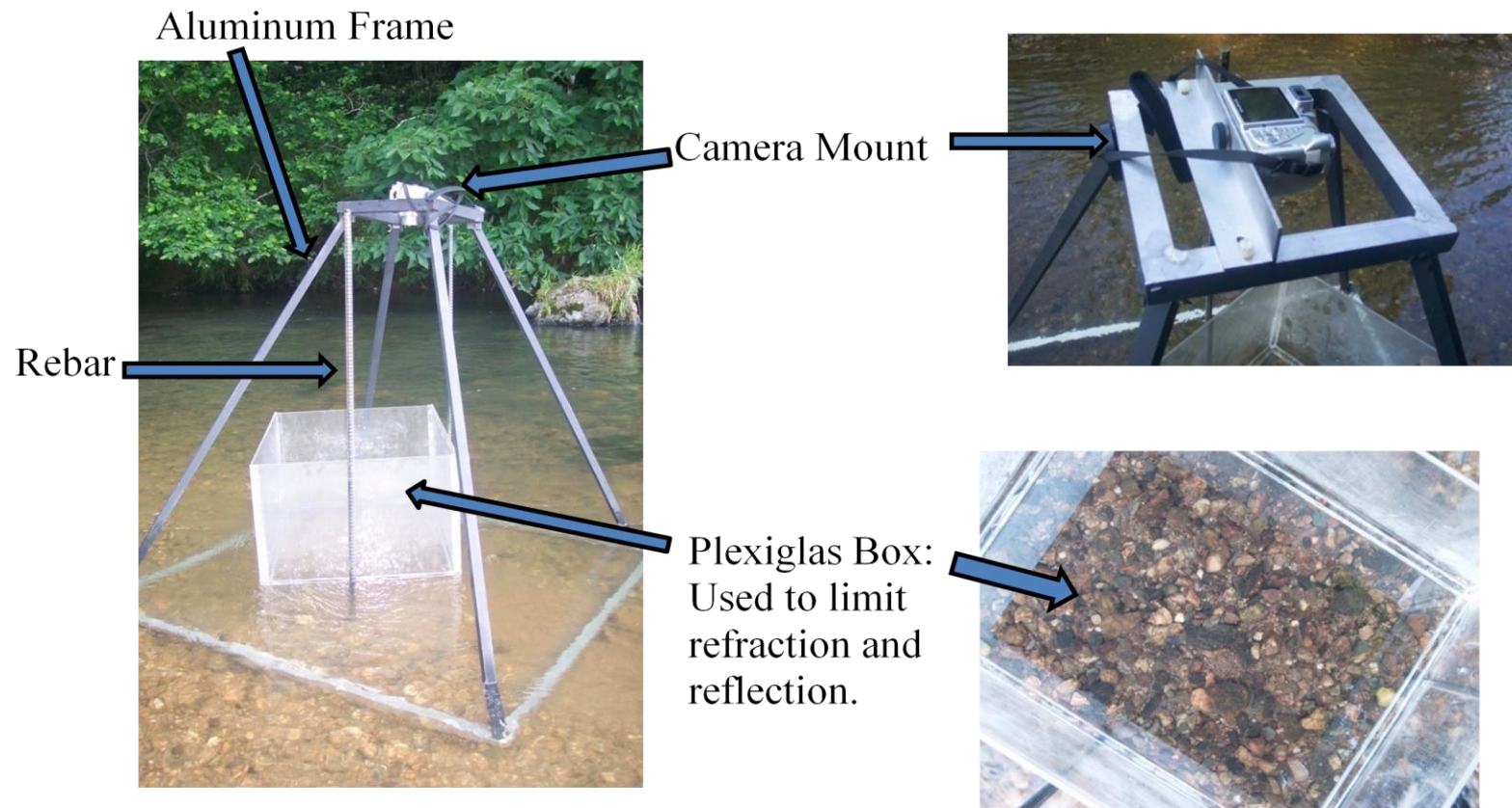


Figure 2. Calibration images used by the autocorrelation algorithm for 200 mm focal length.



An umbrella was also used to further reduce reflection off the Plexiglas box.

Figure 3. The apparatus.

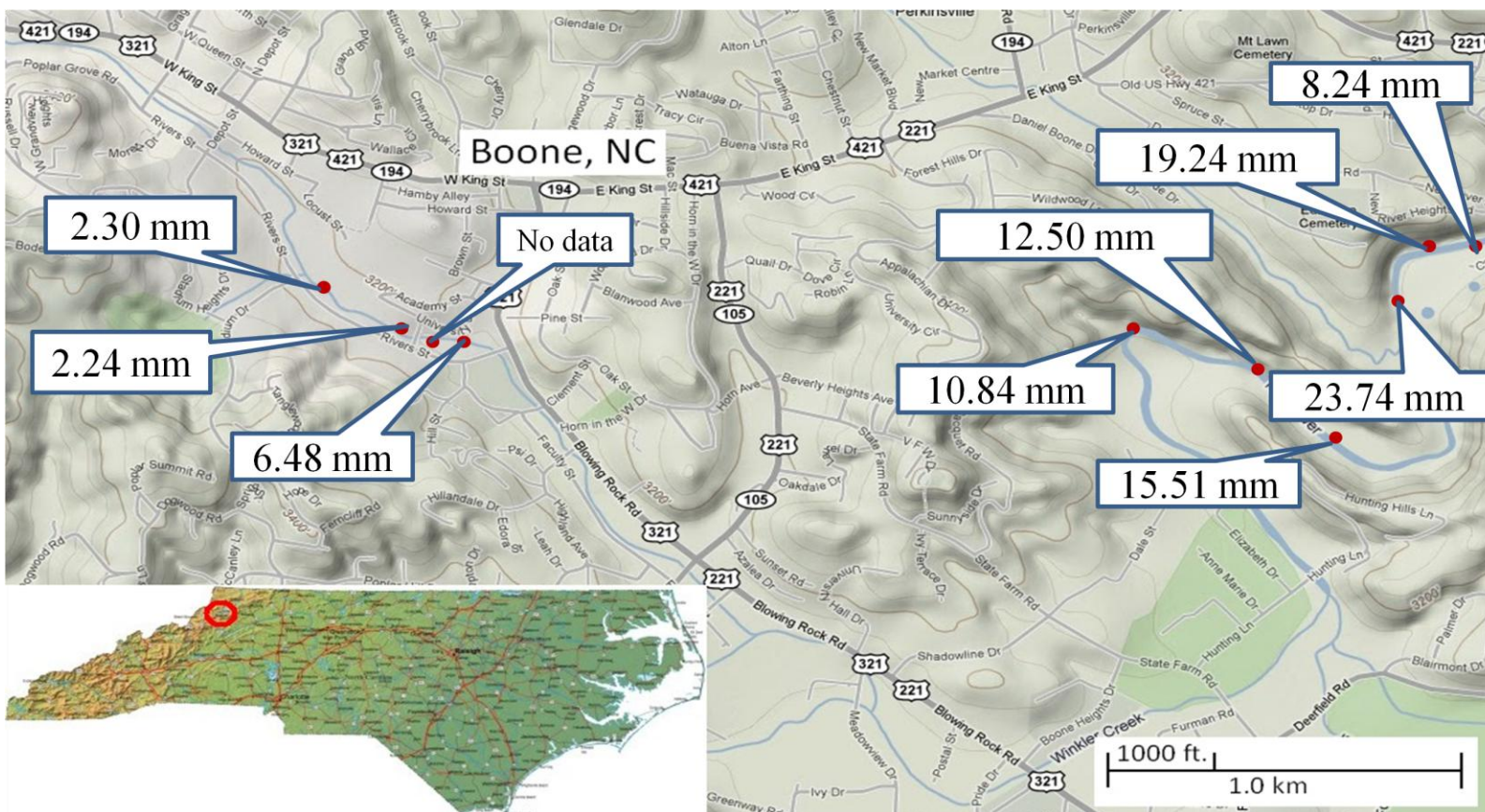


Figure 4. The average predicted mean grain diameter for the sites listed on the map in the methods section. These results were generated from optimized parameters as discussed in the results section.



Sample data images collected from the New River in Summer of 2008. These are a good example of the quality of image necessary for the algorithm to function properly.

BatchID:200mm_500off_1segcal_16segimg
 ImageFile:LB01200mm.tif
 Calibration file:200mm_1Seg_500offset_v70.mat
 Image processing parameters:
 Cal numorientations: 8
 AGS numorientations: 8
 Cal maxoffset: 500
 AGS maxoffset: 500
 Cal num segments: 1
 AGS num segments: 16
 num size bins: 10
 AGS bindist algorithm ID: 1

Segment breakdown by size bin:

0.05000000	0.38000000	0.75000000	1.50000000	3.35000000	6.35000000	12.00000000	20.50000000	37.50000000	75.00000000
0.00000000	0.00000000	0.12055179	0.08484760	0.79460061	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.50536089	0.10413724	0.00000000	0.35193062	0.00000000	0.00000000	0.00000000	0.03857124
0.00000000	0.00000000	0.64310756	0.24317040	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.11372204
0.00000000	0.00000000	0.00000000	0.00000000	0.86707592	0.13292408	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.11027708	0.00000000	0.80963232	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.08009060
0.00000000	0.00000000	0.00000000	0.34656776	0.65343224	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	1.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.12194274	0.15975329	0.53209677	0.00000000	0.00000000	0.00000000	0.18620720
0.00000000	0.00000000	0.00000000	0.00000000	0.29581251	0.49659039	0.00000000	0.04060772	0.16698937	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.68584297	0.00000000	0.31415703	0.00000000
0.00000000	0.00385151	0.55126197	0.00000000	0.00000000	0.17507972	0.26980680	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.02351737	0.00000000	0.97648263	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000	0.26556656	0.31004876	0.42438468	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.52905037	0.37775321	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.09319642
0.00000000	0.00000000	0.09526377	0.00000000	0.90473623	0.00000000	0.00000000	0.00000000	0.00000000	0.00000000
0.00000000	0.00000000	0.00000000	0.00000000	0.00000000	0.97115109	0.02884891	0.00000000	0.00000000	0.00000000

Mean diameter: 7.55465620

Standard error: 6.49285671

Figure 5. A sample output from the algorithm. Information unique to this run is listed at the top to the left. The columns of numbers are representative of the grain sizes as described by the calibration set. The rows represent results for each section in the image processed.

Biographical Information

I was born in West Palm Beach, Florida September 21, 1965. I lived in a variety of locations while growing up and attended many different schools in the following states: California, Maryland, Vermont, New Mexico, and Virginia. My young life was totally involved in the family business of raising and training Arabian horses where I learned about everything from equine veterinary medicine to accounting and marketing. My love of nature and science is strongly rooted in these early years on the farm.

I married in 1989 in Lexington, Va. where two of my three children were born. The next ten years were spent in the family business: owning and operating several different restaurants. After my divorce, I continued working in restaurants and welcomed my third child in Lenoir, NC. At the time I knew I wanted to continue my education, but options were limited for a single parent. It took a couple of years to work out the logistics, but in August of 2002 Appalachian State University accepted my application.

I graduated Magna Cum Laude in December 2007, receiving my Bachelor's Degree in Physics. I entered the Master's Program shortly thereafter for Engineering Physics and received my Master's Degree in May 2010. My aspirations were fully realized when I accepted a position with the environmental science program here at Appalachian State immediately upon graduation. I may at some later date pursue my doctorate in environmental or atmospheric sciences. PhD or not, I will continue to collaborate in research endeavors and to educate and encourage as many of the inquisitive young men and women who are willing to listen!